0727

PB 206 823

TRC THE RESEARCH CORPORATION of NEW ENGLAND

210 Washington Street, Hartford, Connecticut 06106

203 527-4101

SPACE AND TIME VARIABILITY OF SO_2 AND PARTICULATE CONCENTRATIONS IN CONNECTICUT

Prepared by

Norman E. Bowne

September 1971

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

Prepared under Contract CPA 70-155 for Division of Meteorology, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, North Carolina

65

BIBLIOGRAPHIC DATA SHEET	1. Report No. APTD-0929	2.	3. Recipient's Accession No.
4. Title and Subtitle Space and Time Variability of SO_2 and Particulate Concentrations in Connecticut			5. Report Date September 1971 6.
7. Author(s) Norman E. Bowne			8. Performing Organization Rept. No.
9. Performing Organization Name and Address The Research Corporation of New England 210 Washington Street Hartford, Connecticut 06106			10. Project/Task/Work Unit No. 11. Contract/Grant No. CPA 70-155
12. Sponsoring Organization Name and Address Environmental Protection Agency National Environmental Research Center Division of Meteorology Research Triangle Park, North Carolina			13. Type of Report & Period Covered 14.
15. Supplementary Notes			

16. Abstracts The broad objective was to determine the space and time variability of air quality, on time scales two hours or longer and on a minimum space scale of 5000 feet square in urban, suburban and rural locations and to relate this to the use of monitoring and models in tactical and strategic air quality operations. An attempt is made to determine how great a bias might be introduced by improper sampling procedures. The air quality observations were obtained at 30 locations on 25 different sampling days between September 1968 and April 1969. Estimates of SO_2 concentration were made by two models for four selected areas of the state. Observed air quality data were obtained from 2-hourly sulfur dioxide measurements, 2-hourly COH measurements and 24-hourly particulate measurements. Statistical tabulations were prepared for 2-hour and 24-hour sulfur dioxide and particulate concentrations, both observed and estimated by model methods. Frequency distributions were prepared and the observed data were log-normal. Spatial distributions of air quality model estimates were examined in three regions, one of which was actually a double region in the eastern portion of the state for both models. Modeling for both short and long-term concentration estimates can be performed but it is very clear that a good source inventory is required. This is especially important for short-term models.

17. Key Words and Document Analysis. 17a. Descriptors

Air pollution
Sulfur dioxide
Particles
Urban areas
Monitors
Sampling
Atmospheric diffusion
Mathematical models
Statistical analysis
Concentration (composition)

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group 13B

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Unlimited	UNCLASSIFIED	66
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70) USCOMM-DC 40329-P71

TRC THE RESEARCH CORPORATION
of NEW ENGLAND
210 Washington Street, Hartford, Connecticut 06106
203 527-4101

SPACE AND TIME VARIABILITY OF SO_2 AND PARTICULATE CONCENTRATIONS IN CONNECTICUT

Prepared by
Norman E. Bowne

September 1971

Prepared under Contract CPA 70-155 for Division of Meteorology, National Environmental Research Center, Environmental Protection Agency, Research Triangle Park, North Carolina

TABLE OF CONTENTS

Section	Title	Page
1.0	INTRODUCTION	1
2.0	EQUIPMENT AND PROCEDURES	3
3.0	FREQUENCY DISTRIBUTION OF OBSERVED DATA	11
3.1	Sulfur Dioxide	11
3.2	Particulates	20
4.0	VARIATION WITH TIME	27
4.1	SO_2 Data	27
4.2	Particulate Data	35
5.0	SPATIAL VARIABILITY FOR SO_2	42
5.1	Synoptic Variation	42
5.2	Model Calculated Variability on SO_2 Concentrations	46
6.0	CONCLUSIONS	61
	References	63

1.0 INTRODUCTION

The broad objective of this program was to determine the space and time variability of air quality, on time scales two hours or longer and on a minimum space scale of 5000 feet square, in urban, suburban and rural locations and to relate this to the use of monitoring and models in tactical and strategic air quality operations. An attempt is made to determine how great a bias might be introduced by improper sampling procedures.

The air quality observations used included those of SO_2 and particulate concentration obtained at 30 locations on 25 different sampling days during the period September 1968 to April 1969.

Estimates of SO_2 concentration were made by two models for four selected areas of the state. The TRC regional model was used to obtain two-hour estimates in areas nominally known as Bridgeport, New Britain, Norwich and Chaplin. The Martin-Tikvart Model (Air Quality Display Model) was used to prepare average estimates for these same areas using the same source inventory, stability Class D and a wind rose based on the observed data from the 25 sampled days. The limited meteorological sampling period precluded additional delineating of the wind rose data.

Observed air quality data were obtained from 2-hourly sulfur dioxide measurements, 2-hourly COH measurements and 24-hourly particulate measurements. It was noted that some impingers had different flow rates at the conclusion of the observation program than those obtained on initial calibration. Subsequent investigation indicated that the reduced flow was created by particulate collection on the limiting orifice. The investigation also showed the decrease in flow to be linear, therefore a linear correction was applied to the sulfur dioxide data to

account for flow changes.

Two-hour values of observed suspended particulates were obtained by distributing the daily average value from high volume sampler measurements according to the coefficient of haze (COH) data observed for 2-hour periods as described in Section 3.2.

Statistical tabulations were prepared for 2-hour and 24-hour sulfur dioxide and particulate concentrations, both observed and estimated by model methods. Frequency distributions were prepared and the observed data were log-normal.

Spatial distributions of air quality model estimates were examined in three regions, one of which was actually a double region in the eastern portion of the state for both the TRC regional model and the APCO Air Quality Display Model.

Procedures and data used are discussed in Section 2, distribution statistics are presented in Section 3, time variability is illustrated in Section 4 and space variability is shown in Section 5. Conclusions are listed in Section 6.

2.0 EQUIPMENT AND PROCEDURES

Data were collected at 30 fixed stations shown in Figure 1, along highways by mobile samplers and above the ground by an instrumented aircraft. The mobile routes used depended on wind direction. The routes selected on any given day were chosen to be at right angles to the wind.

Research Appliance Corporation sequential samplers with impingers were employed at 27 locations for determination of SO_2 concentration by the West-Gaeke method. Davis SO_2 analyzers were used at Stamford, New Haven and Hartford. Mobile sampling of SO_2 was accomplished with the SIGN-X sampler which is quite similar to the Davis conductiometric device.

The RAC sequential samplers absorbed the SO_2 from the air sample for a 2-hour period in 0.1 N sodium tetrachloromercurate. Nonvolatile dichlorosulfitomercurate ions are formed in this process. Addition of acid-bleached pararosaniline and formaldehyde to the complexion produces red-purple pararosaniline methylsulfonic acid from which concentration may be determined spectrophotometrically.

The Davis and SIGN-X continuous SO_2 analyzers operate by determining the electrical conductivity of a solution of atmospheric gases.

Limiting orifices were employed to insure critical flow for the RAC sequential samplers. As noted previously, they accumulated particulate matter and the flow rate decreased during the data collection period. Six of the samplers were selected and run for five days in the laboratory at Hartford sampling outside air. Flow rates were measured at 8:00 A.M., noon and 4:00 P.M. Flow rates decreased linearly in all cases, although at different rates. On this basis all sequential data were adjusted for linear changes in flow rate based on initial calibration obtained prior to placing

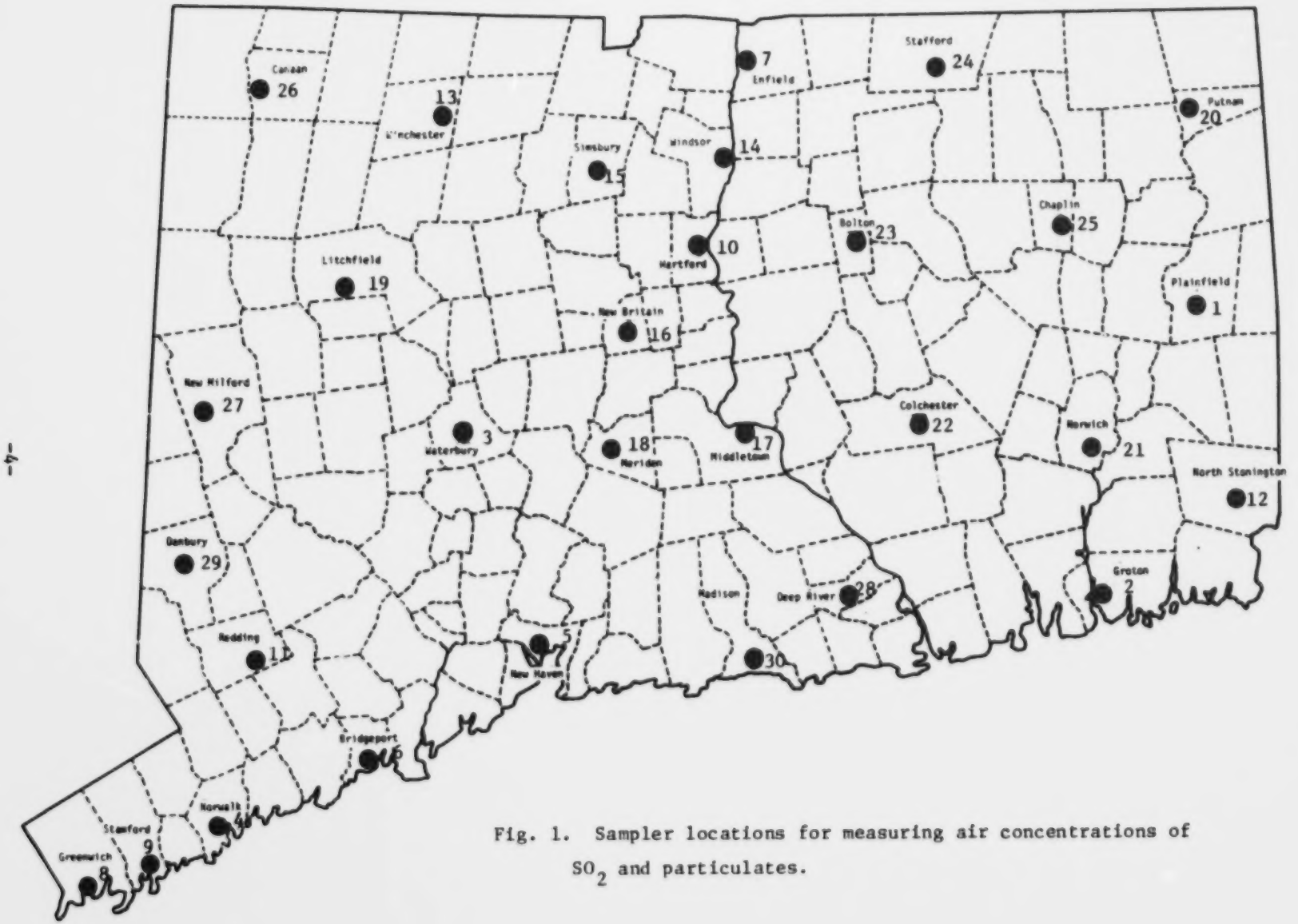

Fig. 1. Sampler locations for measuring air concentrations of SO_2 and particulates.

them in the field and final calibration obtained at the end of the sampling program. The Davis instruments were operated by municipal agencies and calibrated by the State Health Department. They did not appear to have changes in flow rates according to the State laboratory. The SIGN-X samplers were monitored in use for flow rates read directly from the rotometer on the device. The rotometers were calibrated against a standard flow device.

Particulate weight was obtained by exposing filter paper for 24-hour periods in standard high volume sampling equipment. Filters were weighed before and after sampling, flow rates were measured before and after sampling and the total weight of material on the 8" x 10" filter was converted to mass concentration.

COH values were obtained from paper tape filter samplers. The paper tape samplers operated simultaneously with the SO_2 sequentializers to collect twelve 2-hour samples of particles. These samples were analyzed by measuring the amount of light absorbed by the particles using a Gelman densitometer. The results of this type of analysis are reported in COHS per 1000 lineal feet of air. The coefficient of haze (COH) is a measure of light absorption by particles and is defined as that quantity of light scattering solids producing an optical density [$\log (I_o/I_T)$] of 0.01 when measured by light transmittance. One COH is equivalent to 97.73% transmission of light.

Sampling locations were selected to represent urban, suburban and rural classification of land use. Table 1 lists the locations and classification. The four areas selected for detailed analysis include two urban, one suburban and one rural area. The urban areas were centered on Bridgeport and New Britain, the suburban area was centered on Norwich in the eastern part of the state and was joined on the north by the rural area centered on Chaplin.

TABLE I

Sampling Locations

STATION	TOWN	CLASSIFICATION
1	Plainfield	Urban
2	Groton	Suburban
3	Waterbury	Urban
4	Norwalk	Urban
5	New Haven	Urban
6	Bridgeport	Urban
7	Thompsonville	Urban
8	Greenwich	Urban
9	Stamford	Urban
10	Hartford	Urban
11	Redding	Rural
12	North Stonington	Rural
13	Winsted	Urban
14	Windsor	Suburban
15	Simsbury	Suburban
16	New Britain	Urban
17	Middletown	Urban
18	Meriden	Urban
19	Morris Dam	Rural
20	Putnam	Urban
21	Norwich	Suburban
22	Colchester	Suburban
23	Bolton	Rural
24	Stafford	Rural
25	Chaplin	Rural
26	Falls Village	Rural
27	New Milford	Rural
28	Deep River	Rural
29	Danbury	Urban
30	Madison	Suburban

The distribution of sources averaged over 10,000 feet grid squares is shown in Figures 2 to 4 for the four areas selected for analysis. This was the source data used for input to the Air Quality Display Model Calculations. Data on a 5000 foot square separation were used for the TRC model. It was not felt that such detail would add any information to the AQDM calculations.

The starred grid point is the center of the nine contiguous squares used for calculation and comparison described in Section 5. Point source positions are identified by the letter P.

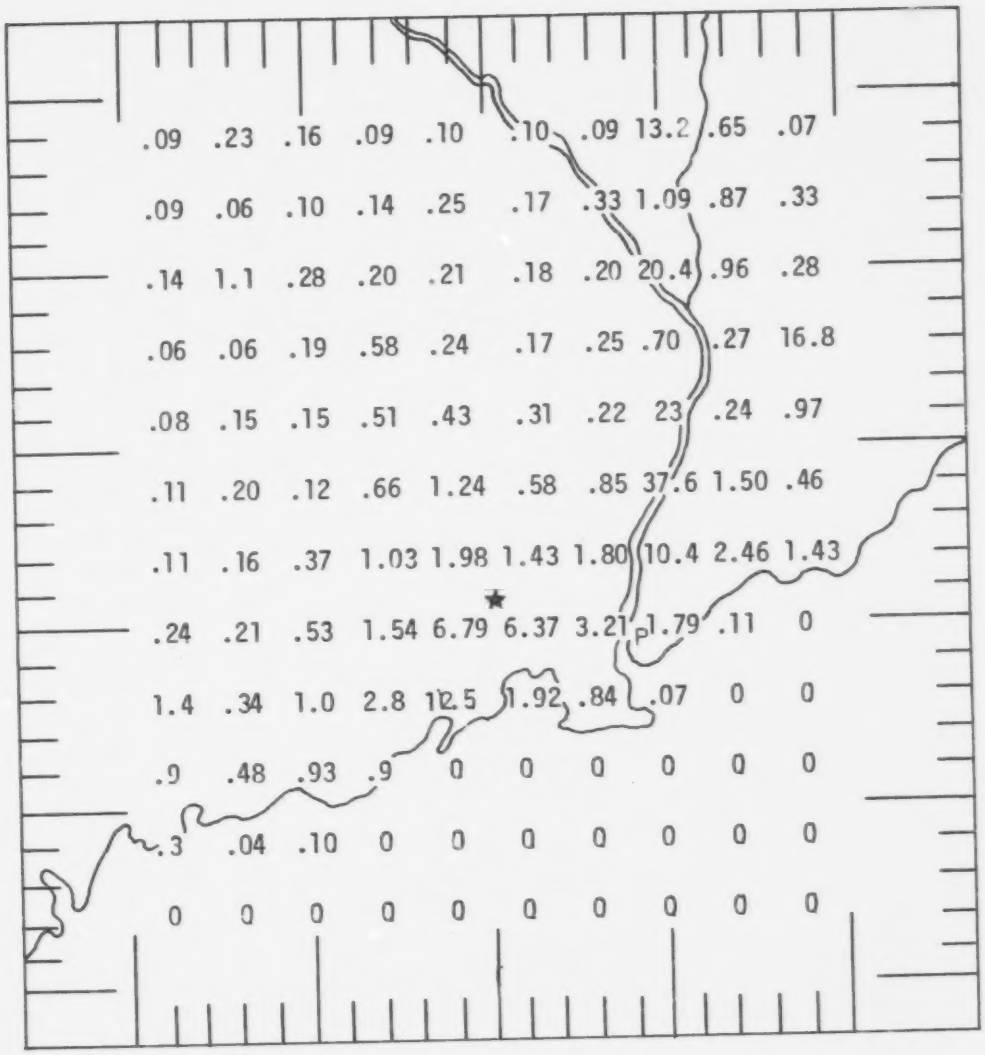

Fig. 2. SO_2 source terms by 10,000 ft grid squares tons/day. Bridgeport sector. P represents point source locations, a ★ is the center of the prediction area grid.

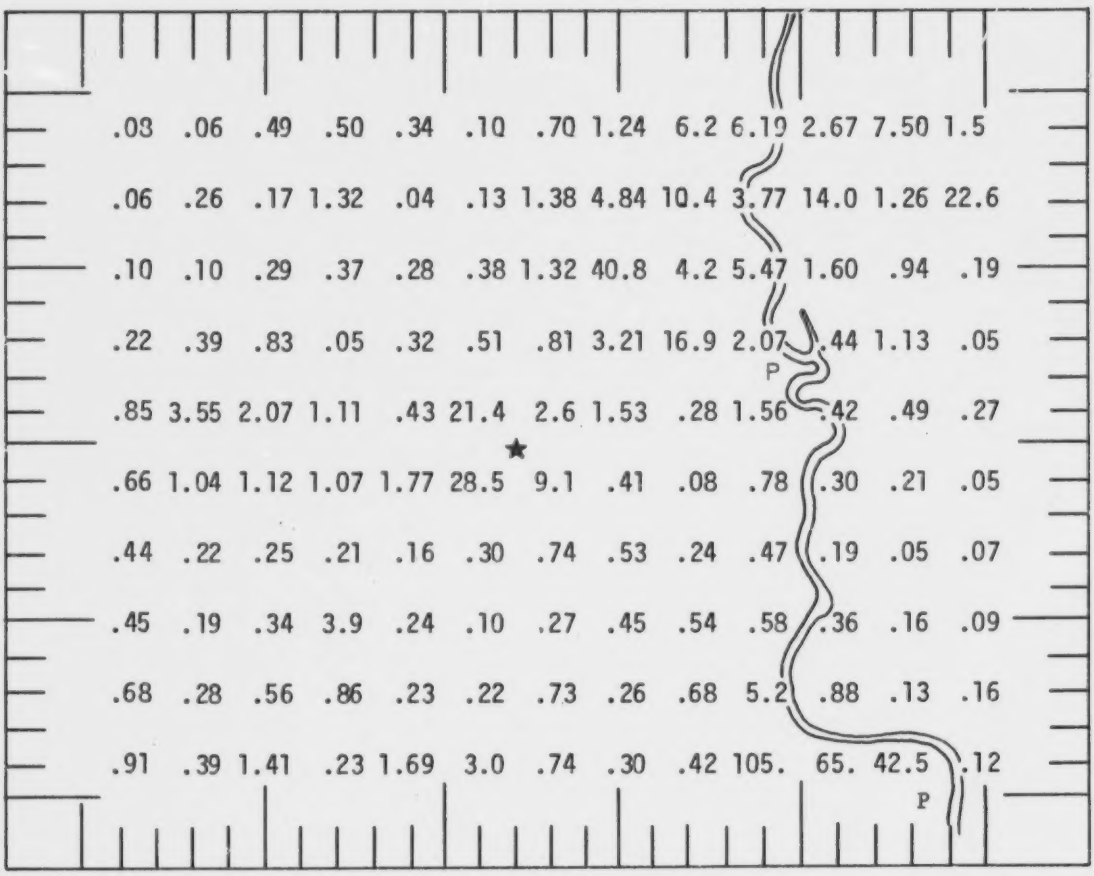

Fig. 3. SO_2 source terms by 10,000 ft grid squares tons/day. New Britain sector. P represents a point source, a ★ is the center of the prediction area grid.

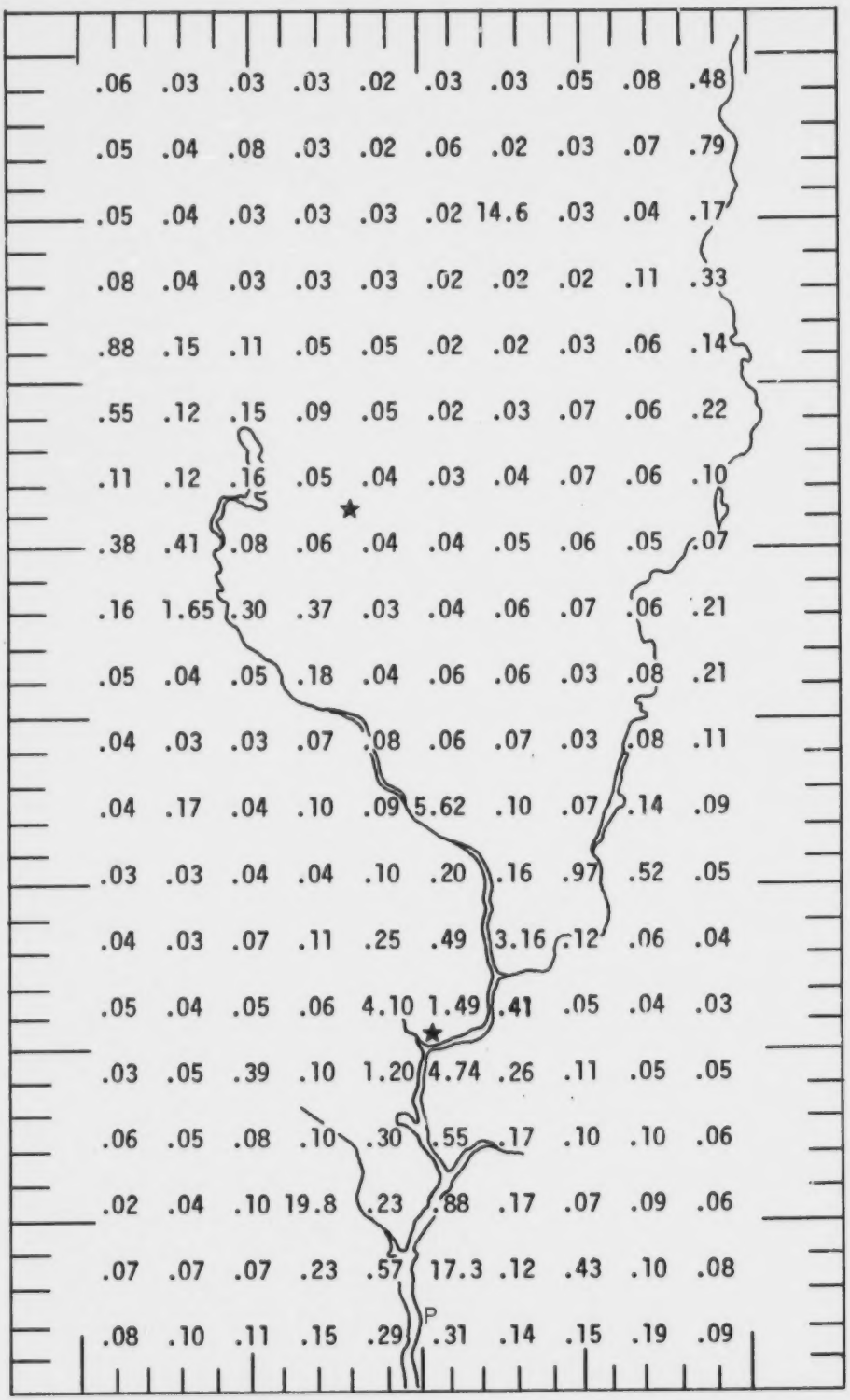

Fig. 4. SO_2 source terms by 10,000 ft grid squares tons/day. Chaplin (upper), and Norwich (lower) sectors.

3.0 DISTRIBUTION OF OBSERVED DATA

3.1 Sulfur Dioxide

Subsequent recalibration of equipment after the initial verification report was written, Hilst (1969), indicated that flow rates had decreased from their initial values. Tests made on the equipment indicated that these rates decreased in a linear fashion. The assumption was made that the flow rates decreased during the data collection period in a linear fashion and the concentration values were corrected according to linear flow calibration curves established for each instrument. This resulted in raising the grand average of all observed SO_2 concentrations from 52.5 $\mu g/m^3$ (.02 ppm) reported originally to 76 $\mu g/m^3$ (.029 ppm). Mobile sampled data were not used to obtain distribution statistics because absolute calibration of the equipment was difficult. Relative concentrations appear to be valid.

The frequency distribution of all observed two-hour averaged SO_2 data is shown in Figure 5. Data from all stations are grouped together on this diagram so this represents the distribution of all observations regardless of location or test. TRC model estimates are indicated also. The observed data are clearly log-normal, the estimated values are not. The slope of the observed line is steeper than those normally associated with urban sampling points. Larsen (1969) found that the standard geometric deviation, s_g, was near 2 for SO_2 rather than 3, as found in Figure 5. The mixture of urban and rural sampling sites, leading to more low concentrations, accounts for the steeper slope. Model estimates result in an overall under-prediction of sulfur dioxide concentrations. The slope appears reasonable above the 70%

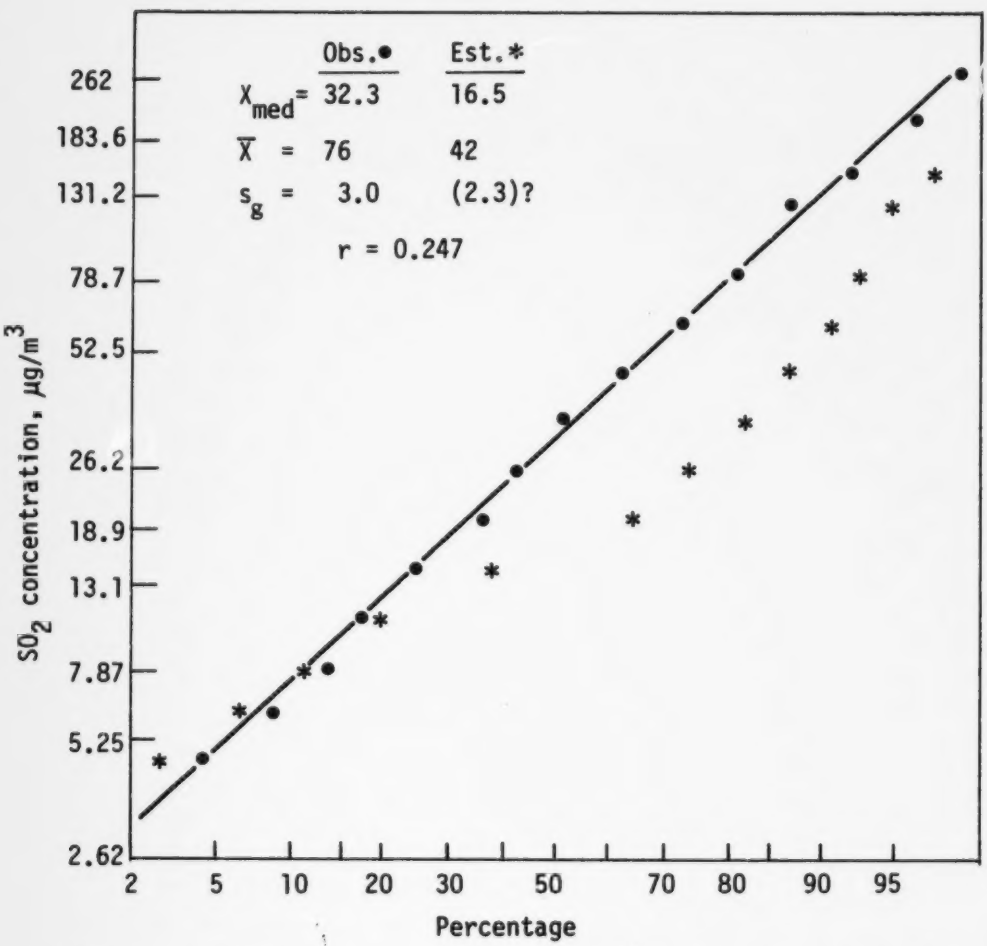

Fig. 5. Connecticut observed SO_2 data; all stations, 2 hour averages, Connecticut estimated SO_2 data; all stations, 2 hour averages.

level but low concentrations which agree well with observed data create a flat section through the middle of the distribution. The increased percentages of estimates at low levels is created by the addition of background, or in the case of the TRC model, of border flux of material. When added appropriately for measured flux of significance such as that from New York, it is highly important, but when added as an average value from areas of minor sources it is apparently quite incorrect.

Data averaged for 24-hour periods are shown in Figure 6. Again, the observed data are log-normal but the slope is less than expected because the extremes of the frequency distribution have been lost to the increased averaging period. Once again the estimated values are not log-normal and significantly under-predict the observed distribution.

Joint frequency distributions are shown in Tables 2 and 3 for 2-hour and 24-hour averaged concentrations. Category limits were set logarithmically to provide better separation of small concentration values. The same problem with estimates is apparent as shown in Figure 5 and 6, i.e., there are too many values in the range of 10.6 to 26 $\mu g/m^3$. Table 4 lists other pertinent statistics of the distributions.

The logarithmic standard deviation of the estimated data should be viewed with caution because the distribution of the estimated data was not log-normal. Student's test for the 2-hour means yields a statistic of .342 for 6000 degrees of freedom against the tabulated 90 percent value of 1.28 which indicates the logarithmic means of the observed and estimated distributions are not significantly different. However, this test is rather weak because the distribution of the estimates is not log-normal.

The Kolmogorov-Smirnov two sample test (Siegal, 1956) was employed for

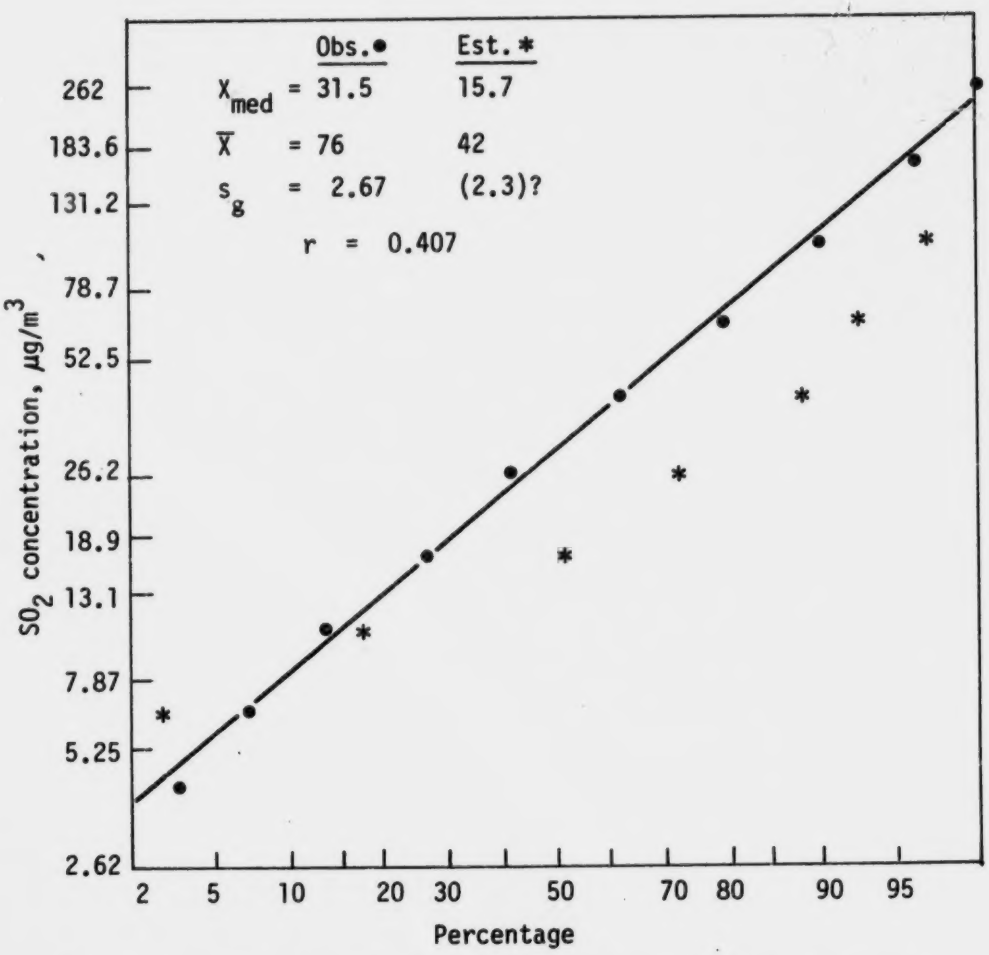

Fig. 6. Connecticut SO_2 data; all stations, 24 hour averages.

TABLE 2

SO_2 Two-Hour Averages $\mu g/m^3$

Observed \ Estimated	4.2	4.2 6.6	6.7 10.5	10.6 16.5	16.6 26	27 42	43 66	67 105	106 165	166 260	261 420	>420	Total
4.2	0	0	0	0	0	0	0	0	0	0	0	0	0
4.2-6.6	0	17	9	102	100	25	5	3	1	4	0	1	267
6.7 - 10.5	0	5	10	74	119	24	9	4	3	5	0	0	253
10.6 - 16.5	0	27	27	211	309	78	56	21	19	12	0	0	760
16.6 - 26	0	27	34	225	372	113	53	22	17	8	3	0	874
27 - 42	0	14	24	147	257	77	42	19	9	11	8	1	609
43 - 66	0	20	38	231	359	163	110	37	31	23	5	4	1021
67 - 105	0	16	35	147	267	134	107	49	38	23	7	3	826
106-165	0	21	10	85	168	136	114	65	32	31	14	14	690
166-260	0	13	5	62	105	50	47	26	25	26	13	7	379
261-420	0	2	4	26	28	15	32	14	13	11	6	17	168
420-660	0	5	2	22	13	2	14	4	8	7	3	9	89
> 660	0	3	1	16	14	2	4	3	2	1	4	6	56
Total	0	170	199	1348	2111	819	593	267	198	162	63	62	5992

TABLE 3

SO_2 Twenty-four-Hour Averages $\mu g/m^3$

Observed \ Estimated	4.2	4.2 6.6	6.6 10.5	10.5 16.5	16.5 26	26 42	42 66	66 105	105 165	165 260	260 420	420 660	Total
4.2	0	0	0	0	5	1	0	0	1	0	0	0	7
4.2-6.6	0	1	0	1	7	0	2	0	1	0	0	0	12
6.6 - 10.5	0	0	1	5	9	3	2	0	0	0	0	0	20
10.5 - 16.5	1	0	1	10	17	5	2	0	0	0	0	0	36
16.5 - 26	1	0	1	11	40	15	6	3	0	0	0	0	77
26 - 42	0	0	1	19	33	18	12	4	0	0	0	0	87
42 - 66	0	1	2	18	37	25	19	4	2	0	0	0	108
66 - 105	0	0	3	6	22	25	19	9	6	4	0	0	94
105 - 165	0	1	0	5	16	17	11	3	6	1	0	0	60
165-260	0	0	1	5	2	6	9	1	2	5	1	0	32
260-420	0	0	0	1	1	2	1	0	2	2	1	2	12
420-660	0	1	0	0	2	0	1	0	0	1	0	1	6
660-1050	0	0	0	1	2	0	0	0	0	1	0	0	4
Total	2	4	10	82	193	117	84	24	20	14	2	3	555

TABLE 4

Mean and Standard Deviations, Arithmetic and Logarithmic of Observed and Estimated SO_2 Concentrations

$\mu g/m^3$

	2-Hour Data		24-Hour Data	
	Observed	Estimated	Observed	Estimated
Arith. Mean	76	42	76	42
S.D.	118	71	94	53
Log. Mean	1.505 (32)	1.230 (17)	1.505 (32)	1.230 (17)
S.D.	.733	.52	.612	.483
Correlation Coefficient	0.247		0.407	

the cumulative frequency distributions. The Kolmogorov-Smirnov two-sample test is a nonparametric statistical test of whether two independent samples have been drawn from the same population. The two-tailed test is sensitive to any kind of difference in the distributions from which the two samples were drawn - differences in central tendency, in dispersion, in skewness, etc. This two-sample test is concerned with the agreement between two cumulative distributions. If the two samples were in fact drawn from the same population distribution, then the cumulative distributions of both samples may be expected to be fairly close to each other, inasmuch as they both should show only random deviations from the population distribution.

To apply the test, cumulative frequency distributions were prepared for each sample (observed and estimated concentrations) using the same intervals for both distributions. For each interval, we subtract one step function from the other. The test focuses on the largest of these observed deviations.

Let $S_{n1}(x) = K/n1$, where K is the number of observations less than or equal to x and n1 is the total number of observations in sample 1. Then $S_{n2}(x) = K/n2$ and

$$D = \text{maximum}\left|S_{n1}(x) - S_{n2}(x)\right| \quad .$$

Critical values of D are listed in Table 5 for three significance levels as a function of sample size. Note that the sample sizes need not be equal.

The 2-hour data on Figure 5 yield D = 0.32 and the 24-hour data plotted on Figure 6 yield D = 0.31. There were 5992 cases for 2-hour data and 555 cases for 24-hour data. The associated critical values of D at the .05 level were 0.025 and 0.0816. The differences in the distributions of the observed and estimated SO_2 concentrations are clearly significant and we cannot consider them to come from the same population.

TABLE 5

Critical Values of D in the Kolmogorov-Smirnov Two Sample Test (Large Samples: Two-tailed Test)

Significance Level	D (Critical)
.10	$1.22\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$
.05	$1.36\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$
.01	$1.63\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$

Doubling the estimates of SO_2 would bring about significantly better agreement except at the low concentration values. Inability of the model to estimate very low values may be caused by the addition of a boundary flux term. A boundary flux was clearly needed for cases with southwest winds, but is of doubtful value unless winds blow from an urban area.

3.2 Particulates

High volume samples and COH data were available from 22 stations to provide enough cases to analyze. Two-hour values of observed suspended particulates were obtained by distributing the daily average value from the high-volume sampler according to the COH data for two-hour periods, i.e.,

$$C_i = C_T \frac{COH_i}{COH_m} \quad (3.1)$$

where C_i is the two-hour particulate concentration for the i'th period, C_T is the average concentration for the day, COH_i is the COH value for the i'th period and COH_m is the mean value for the day.

The frequency distribution resulting from this method of distributing particulate concentrations is shown in Table 6 and Figure 7. TRC model estimates are shown also. The distribution for 24-hour periods from the observed high-volume sampler data is shown in Table 7 and Figure 8. Two-hour averages are log-normal, 24-hour averages are log-normal over most of the distributions but depart from normality in the low and high frequency sections. The departures are not significant from the line drawn on Figure 8 according to the Kolmogorov-Smirnov test.

TABLE 6

Two-Hour Particulate Averages $\mu g/m^3$

Observed \ Estimated	<20	20 39	40 59	60 79	80 99	100 119	120 139	140 159	160 179	180 199	200 219	>219	Total
<20	0	0	6	53	179	63	21	7	2	2	0	0	333
20 - 39	2	7	32	183	507	161	50	28	13	4	2	2	992
40 - 59	1	6	28	124	364	143	48	21	10	3	2	4	754
60 - 79	0	5	19	57	206	70	15	8	5	0	2	5	392
80 - 99	2	3	12	33	115	39	8	9	4	2	0	2	229
100 - 119	0	2	5	37	64	25	11	5	4	1	1	0	155
120 - 139	3	0	10	31	49	19	11	6	1	0	0	2	132
140 - 159	0	3	11	23	28	5	6	3	1	3	0	0	83
160 - 179	0	2	6	11	29	7	5	2	0	0	2	0	64
180 - 199	0	0	5	8	10	8	1	0	1	1	1	0	35
200 - 219	0	0	3	5	9	3	1	3	1	1	0	0	26
229 - 239	0	0	0	6	8	2	0	1	0	0	0	0	17
>239	0	1	6	14	13	3	0	1	0	0	0	1	40
Total	8	29	143	585	1581	548	177	94	42	18	10	17	3252

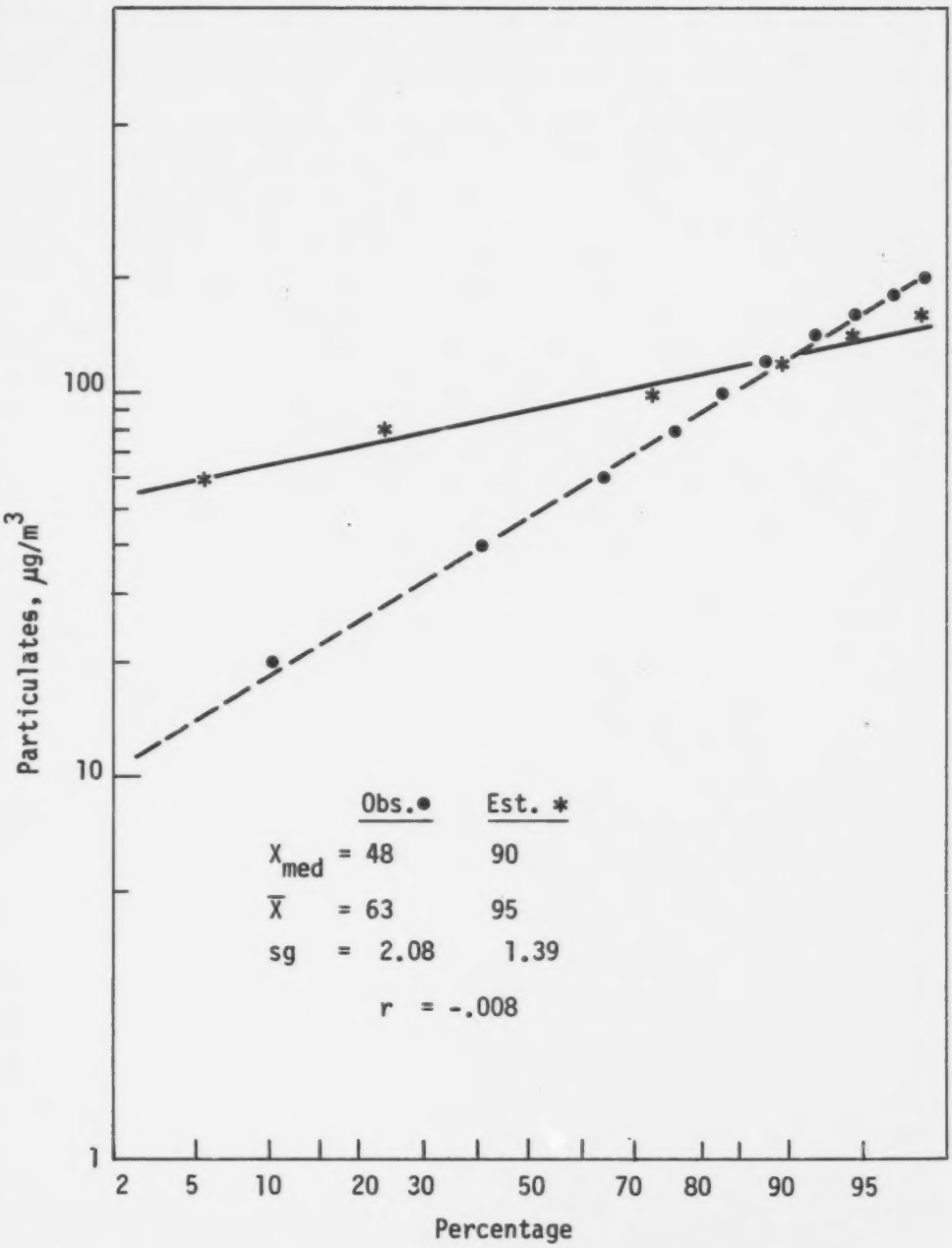

Fig. 7. Connecticut particulate data; all stations, 2 hour average concentration.

TABLE 7

Twenty-four-Hour Particulate Averages $\mu g/m^3$

Observed \ Estimated	<20	20 39	40 59	60 79	80 99	100 119	120 139	140 159	160 179	180 199	200 219	>219	Total
<20	0	0	0	2	5	2	0	0	0	0	0	0	9
20 - 39	0	0	0	16	58	15	4	2	0	0	0	0	95
40 - 59	0	1	0	18	38	6	3	4	1	0	1	0	72
60 - 79	0	0	0	5	24	4	1	0	1	0	0	1	36
80 - 99	0	0	0	2	7	2	2	0	0	0	0	0	13
100 - 119	0	0	0	5	9	2	1	0	1	0	0	0	18
120 - 139	0	0	1	2	2	2	1	0	0	0	0	0	8
140 - 159	0	0	0	5	1	1	0	0	0	0	0	1	8
160 - 179	0	0	0	3	4	0	1	0	0	0	0	0	8
180 - 199	0	0	0	1	0	0	0	0	0	0	0	0	1
200 - 219	0	0	0	0	0	0	0	0	0	0	0	0	0
229 - 239	0	0	0	0	0	1	0	0	0	0	0	0	1
>239	0	0	0	1	2	0	0	0	0	0	0	0	3
Total	0	1	1	60	150	35	13	6	3	0	1	2	272

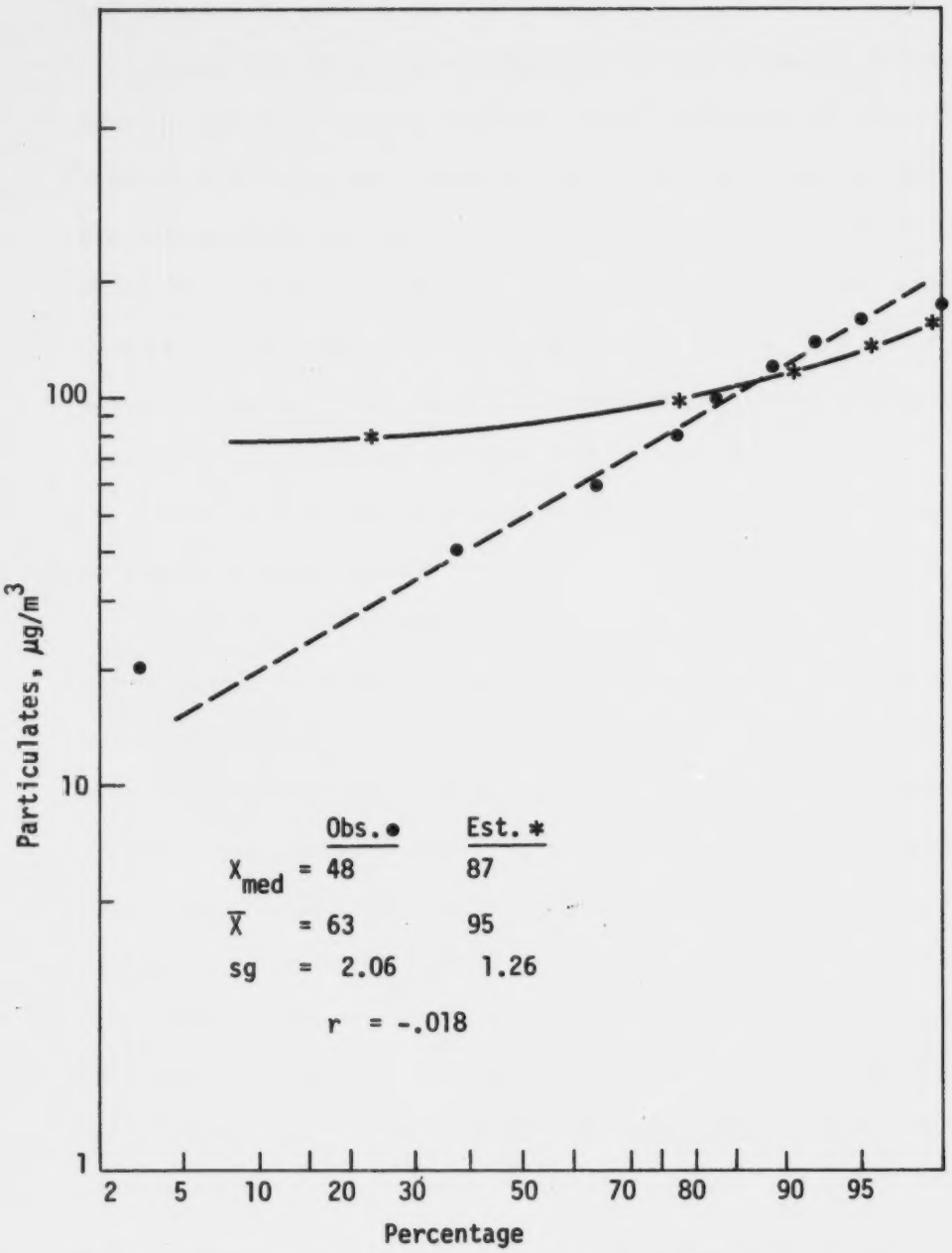

Fig. 8. Connecticut particulate data; all stations, 24 hour average concentration.

The use of COH valves to estimate two-hour values of particulate concentration did not create any change in distribution statistics (mean, median, and standard deviation) in this case.

The effect of adding a background to the estimated concentrations is readily apparent in these figures. Model estimates of particulate concentration were quite low. This was to be expected from the manner in which the source inventory was obtained. In the rural sections of Connecticut where the source inventory was essentially negligible the natural background of particulate concentration was near 40 micrograms per cubic meter. The method of adding this background to estimated values to account for an invalid source inventory is clearly a poor approach.

Table 8 lists the means and standard deviations of the data, as well as the correlation coefficient of observed and estimated data.

Inspection of Figures 7 and 8 clearly indicates the estimated and observed particulate concentrations did not come from the same population. Model predictions did not show the appropriate range of concentrations indicated by observations. The cumulative frequency distribution of model estimates yielded both means and medians that were high compared to observed data. If the arbitrary background data were subtracted from the estimated distributions shown on Figures 7 and 8 median concentration values would be comparable, model estimates on the low frequency side of the median would be close to observed data, but the model estimates of high concentrations would be too low. It appears that uniform addition of an arbitrary background to account for sources not included in the model inventory is a poor choice.

TABLE 8

Mean and Standard Deviations, Arithmetic and Logarithmic, of Observed and Estimated Particulate Concentrations

$\mu g/m^3$

	2-Hour Data		24-Hour Data	
	Observed	Estimated	Observed	Estimated
Arith. Mean	63	95	63	95
S.D.	53	37	45	25
Log. Mean	1.711 (51)	1.965 (92)	1.711 (51)	1.965 (92)
S.D.	.335	.124	.258	.097
Correlation Coefficient	-0.008		-.018	

4.0 VARIATION WITH TIME

4.1 SO_2 Data

Diurnal variability of emissions and observed levels of SO_2 are shown in Figure 9. The emission curve is typical of heating load rather than manufacturing or processing. The emission curve is based on heating demand for natural gas provided by the Connecticut Natural Gas Company for residential areas.

Dividing the sampling stations into three categories of land use leads to the urban curve being significantly above the levels of suburban or rural observations. The same general pattern is evident on all curves. A peak is observed about 0900 corresponding to the maximum emissions; a second peak is found in the late evening even though the emissions are decreasing but can be accounted for by increased atmospheric stability. The trend in the observed concentration curves is the same for all locations. A gradual increase in concentration is observed from midnight to 0900. It is not as steep as the increase in emissions because atmospheric stability was decreasing after 0700 on the test days. A general decrease in concentration is noted from 0900 until about 1700 when the minimum occurs. This corresponds to a decrease and leveling off of the emissions curve and decreasing atmospheric stability. After 1700 the emissions curve continues to decrease but the trend in concentration is upward. At this time there was an increase in atmospheric stability on most test days. Apparently the stability increased more rapidly than the emissions decreased at this time. The end points on the curves, 0100 and 2300, are not coincident because the sampling days were

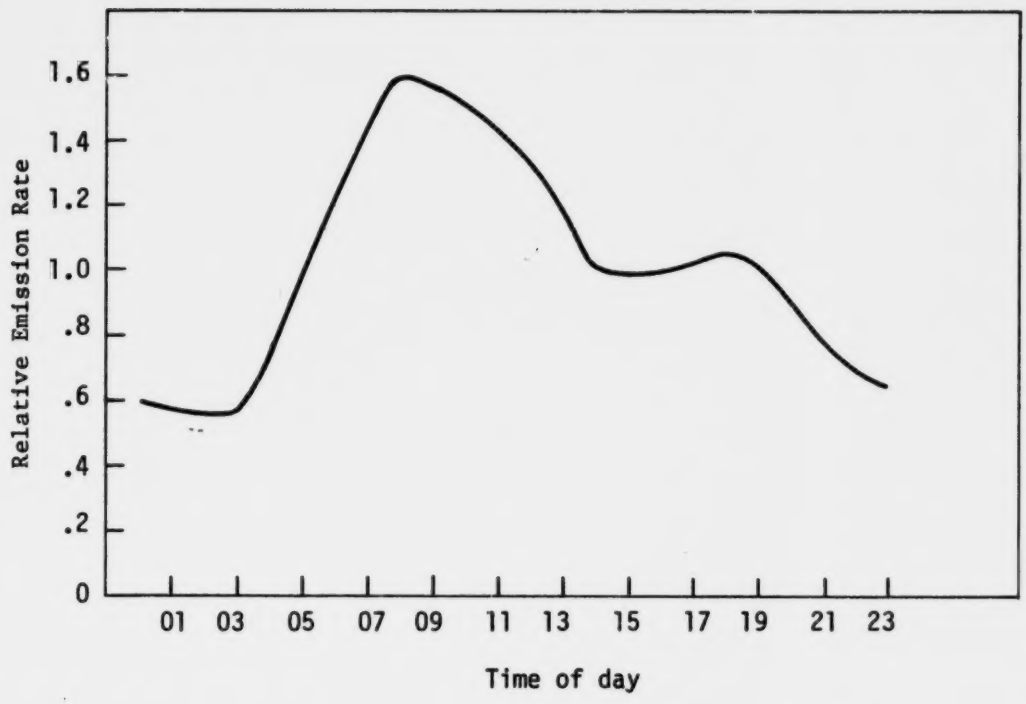

Fig. 9a. Daily variability of SO_2 emission for January.

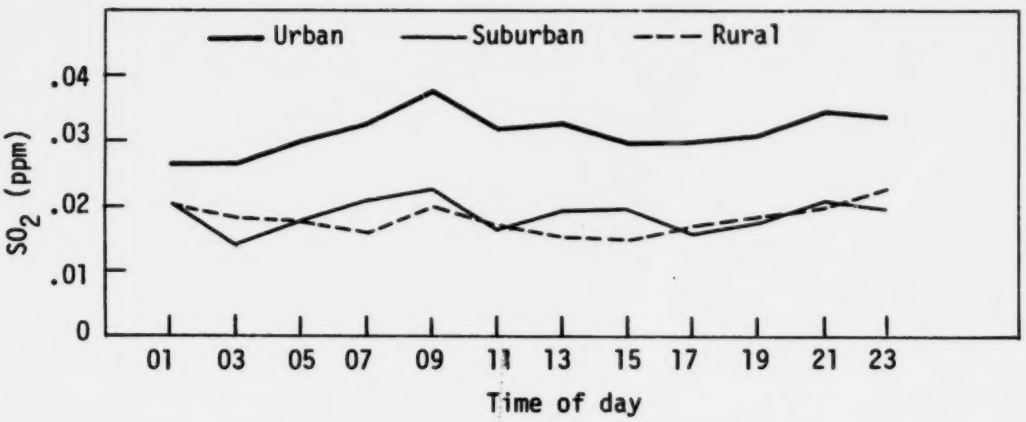

Fig. 9b. Average SO_2 levels for sampling location type.

not contiguous. In most cases only one day per week was sampled.

It is interesting to note that the diurnal emissions curve of Figure 9a has a range factor of 2.9 (i.e., .55 to 1.58) for home heating, while the concentration range factor for observed data in Figure 9b is 1.4 in the urban area and 1.9 in the suburban and rural data. This figure shows that the atmosphere mitigates the diurnal emission effect in the mean. Factory processing would be expected to increase the daytime emissions and decrease these at night for one shift operations. Two or more shifts would yield a flat diurnal distribution but the absence of industry in suburban and rural areas of Connecticut indicates emissions are not the controlling factor. The similarity of the shape of the curves in Figure 9b indicate that the atmosphere dominates and diurnal changes of emission are similar in urban, suburban and rural areas.

The averages at each station are shown in Figures 10 to 13 for all tests. A "D" on the diagram following the station number indicates a Davis instrument was employed rather than a sequentializer analyzed by the West-Gaeke Method. The data were reported in increments of 0.01 ppm (26.2 $\mu g/m^3$). Table 9 reproduces the information of Table 1 to classify the station numbers by land use.

The individual station time averages, shown in Figures 10 to 13, indicate some large diurnal changes in concentration notably at Stamford (9), Hartford (10), Norwich (21) and Madison (30).

Hartford and Stamford were urban sites while Norwich and Madison were suburban. Norwich and Madison were subjected to locally high concentrations from the heating plant at the school where the samplers were located. The erratic diurnal pattern may be typical of poorly selected sampling locations

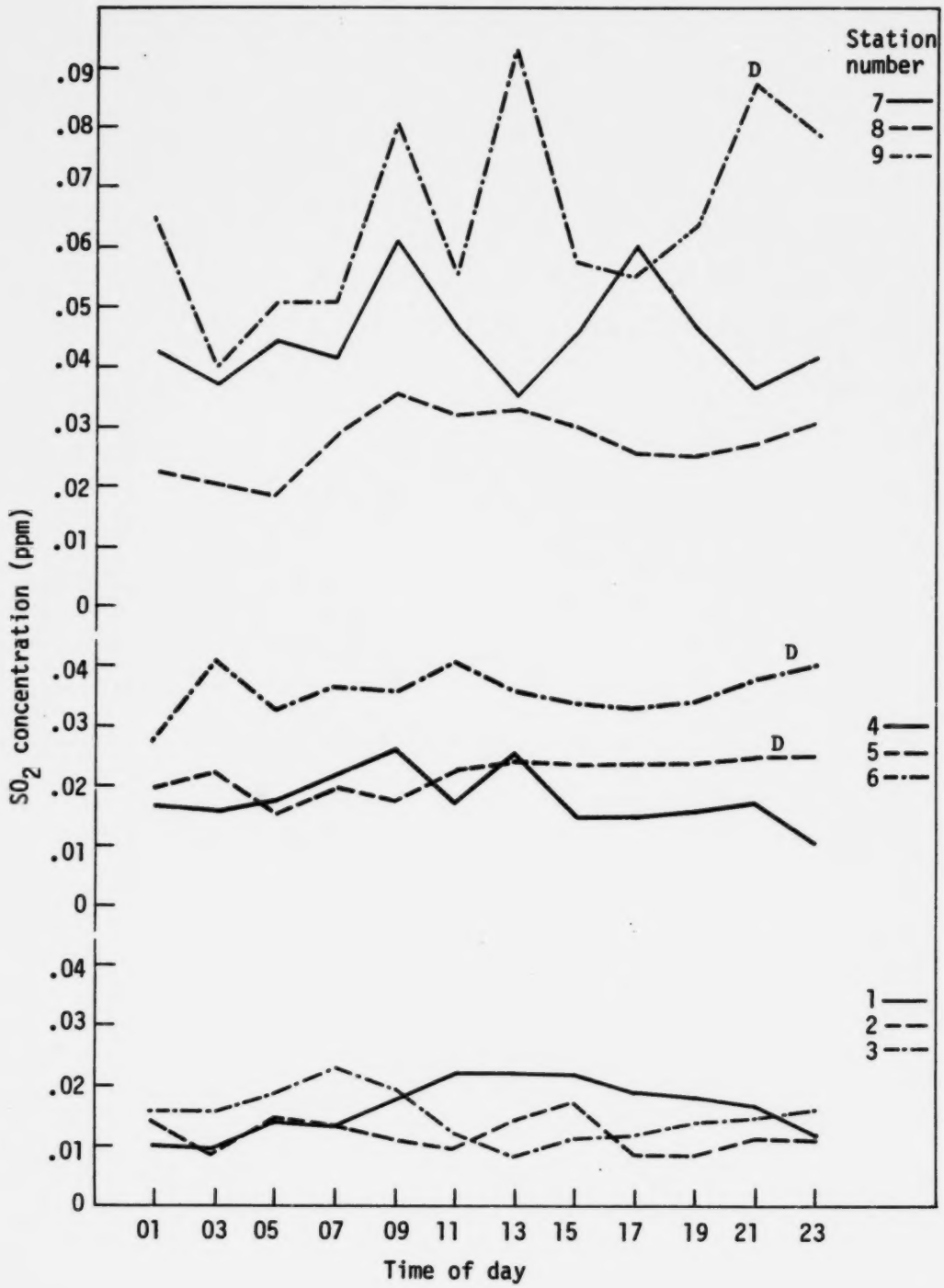

Fig. 10 Average SO_2 concentration by time of day for all sampling days at locations specified.

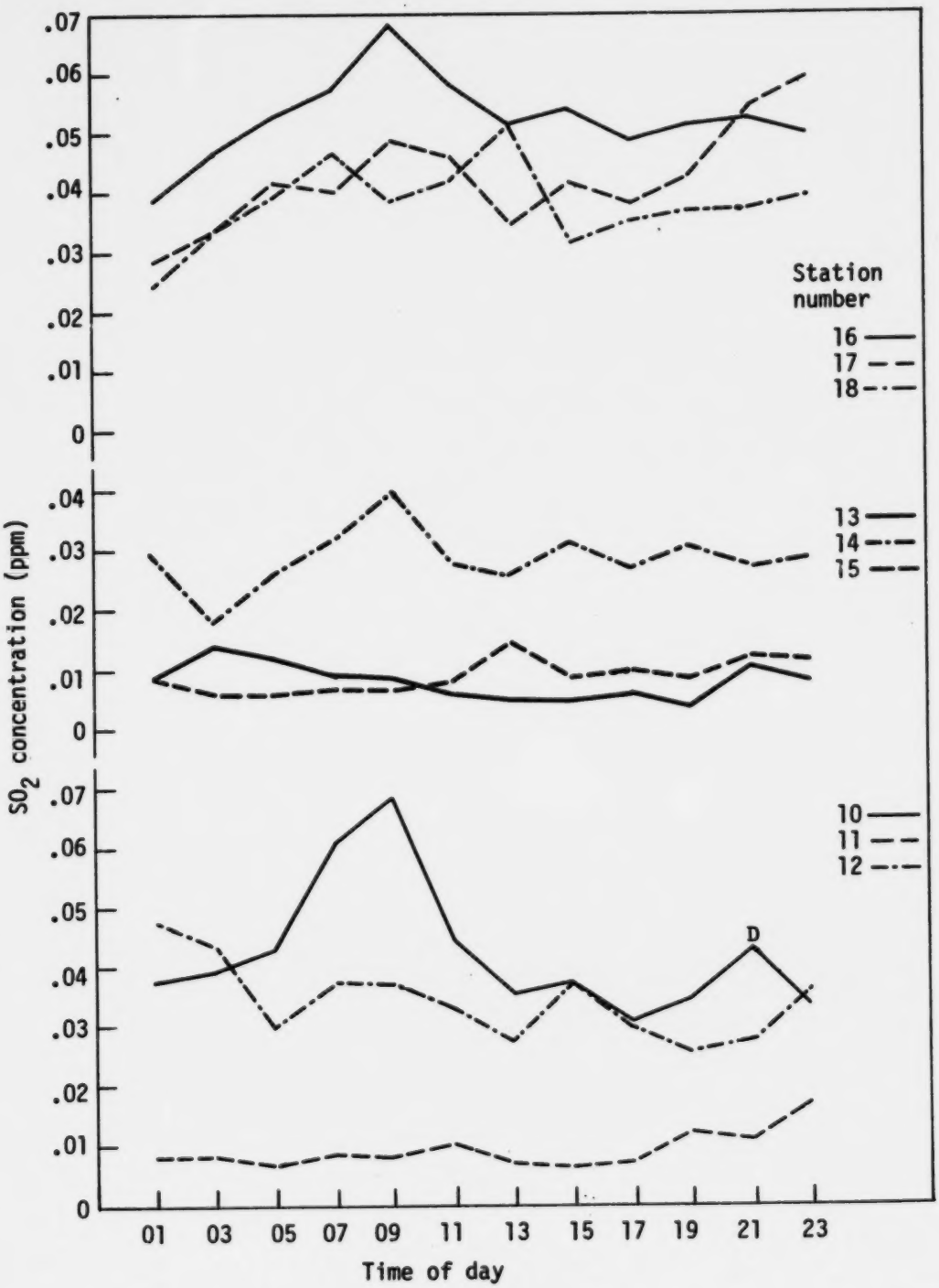

Fig. 11 Average SO_2 concentration by time of day for all sampling days at locations specified.

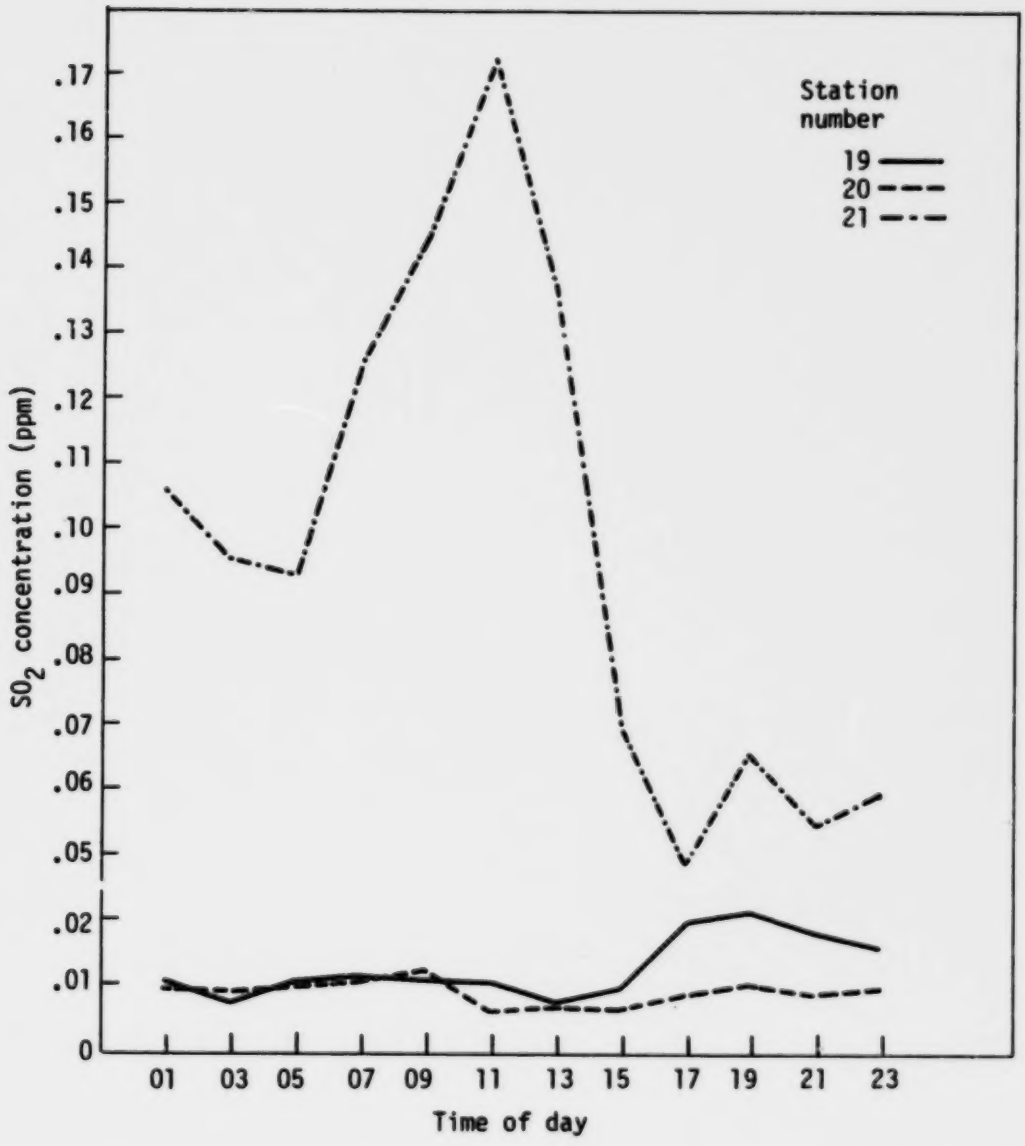

Fig. 12 Average SO_2 concentration by time of day for all sampling days at locations specified.

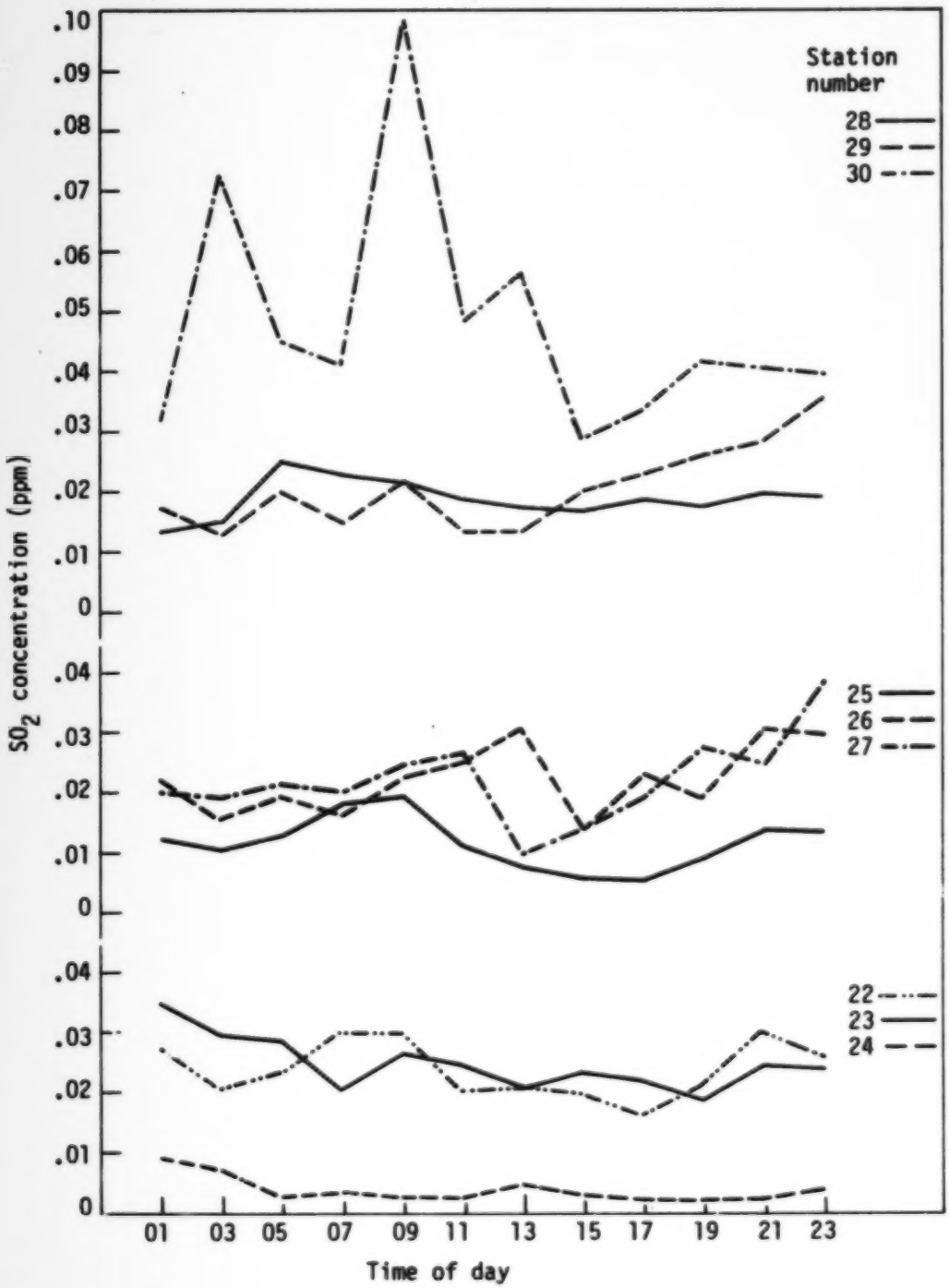

Fig. 13 Average SO_2 concentration by time of day for all sampling days at locations specified.

TABLE 9

Station Classification

Urban		Suburban	Rural
1	9	2	11
3	10	14	12
4	13	15	19
5	16	21	23
6	17	22	24
7	18	30	25
8	20		26
	29		27
			28

subject to local influences.

Stamford data were missing for half the tests so conclusions should not be drawn for that location. Hartford data follow the diurnal trend of emissions quite well. Hartford business is basically one shift so the emissions curve should be reinforced. Middletown and New Britain (16 and 17 on Figure 11) are considered manufacturing towns. The morning peak is indicated at both locations. New Britain (17) shows an exaggerated rise in the evening hours. This is attributed to a combination of third shift operations and changes in dispersion characteristics with time. Very few of the time averaged data show a tendency to meet at the midnight ends of the curve. Data were not collected continuously in this project so the start and ends of diurnal distribution of concentration are not necessarily coincident.

Figure 14 indicates the ratio of concentration change from adjacent two-hour sampling periods. Even though 71 percent of the changes are less than a factor of 2, indications are clear that significant changes of concentration with time do occur.

4.2 Particulate Data

Sufficient data for analysis were available from 21 stations. The diurnal variability is shown in Figure 15 as averages by two-hour period of the particulate concentration. There were 14 urban stations, 4 suburban sites and 3 rural locations available in sufficient quantity (minimum of 10 tests) to qualify for analysis. The general shapes of all the curves are the same exhibiting morning peaks, afternoon lows and secondary evening peaks. The diurnal emission curve for particulates is assumed to be similar to the SO_2 emission curve and the observed concentration curves are certainly

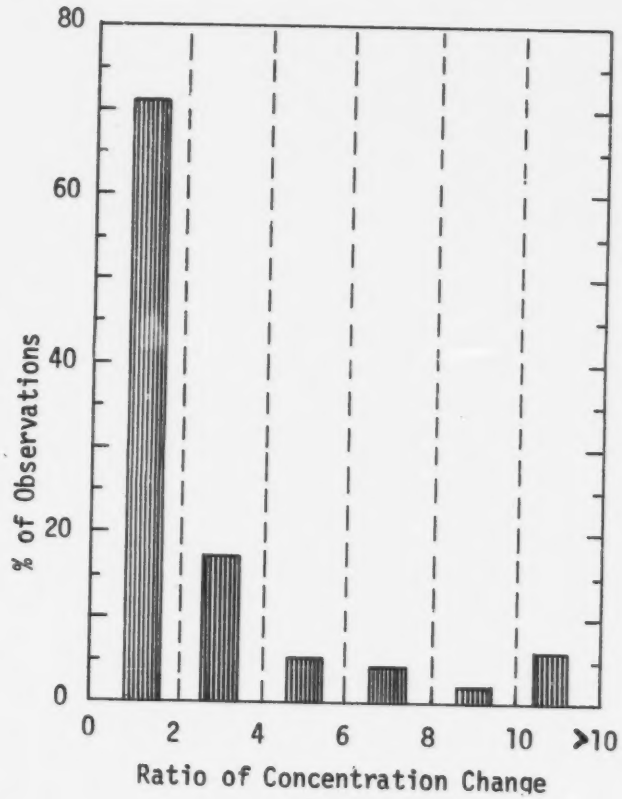

Fig. 14 Two-Hour Change in SO_2 Concentration

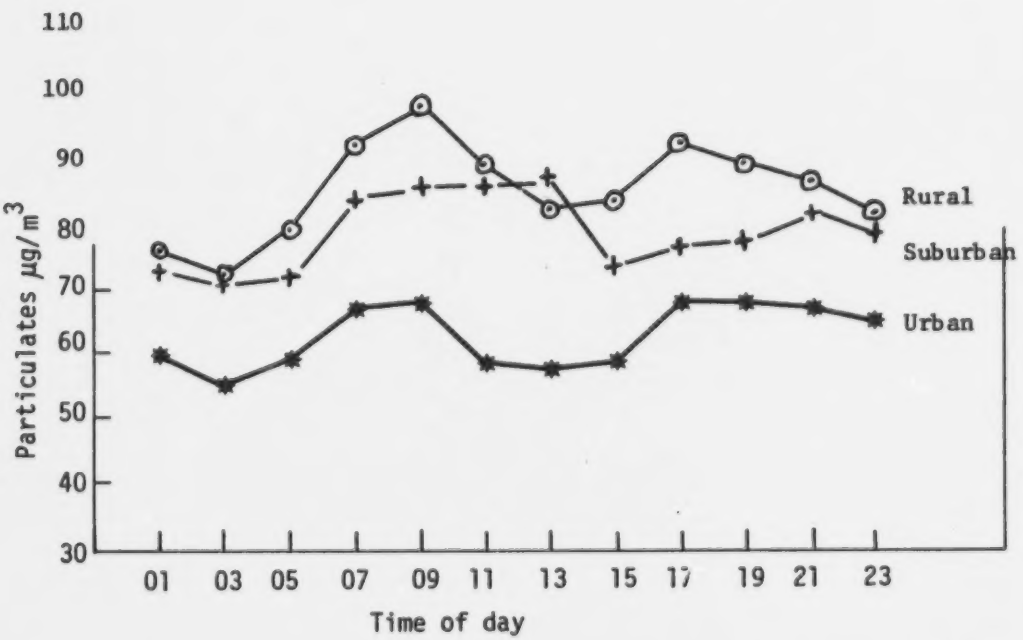

Fig. 15. Average particulate concentration by time of day for 14 urban sites, 4 suburban sites and 3 rural sites in Connecticut.

similar. It may be concluded that the state of the atmosphere plays a predominant role in the variation of concentration during a 24-hour period.

Diurnal trends indicated by the particulate concentration curves are similar to those for sulfur dioxide. Since the observed concentration patterns are similar and the meteorological conditions were the same, we conclude that the emissions change with time of day must be nearly the same as for SO_2.

Rural and suburban averages indicated on Figure 15 suffer from lack of data. It is questionable that average urban concentrations of particulate matter are less than those at rural and suburban locations.

Variation with time at the individual stations is shown in Figures 16, 17 and 18. It should be borne in mind that these data represent soiling or COH changes rather than particulate concentration directly. Individual variation at stations appears larger than for sulfur dioxide. Hartford data (10) on Figure 17 are highly suspect. There were no indications of instrument error but the average concentration level and lack of change with time do not appear normal. Meriden (18) on Figure 18 is suspect for the same reason. The greatest variation was observed at Norwich where the observation site was poorly located. The sampling equipment was on a flat roof in an area of wakes from surrounding buildings. Observed concentrations of both particulates and sulfur dioxide were very high at this site compared to the rest of the locations.

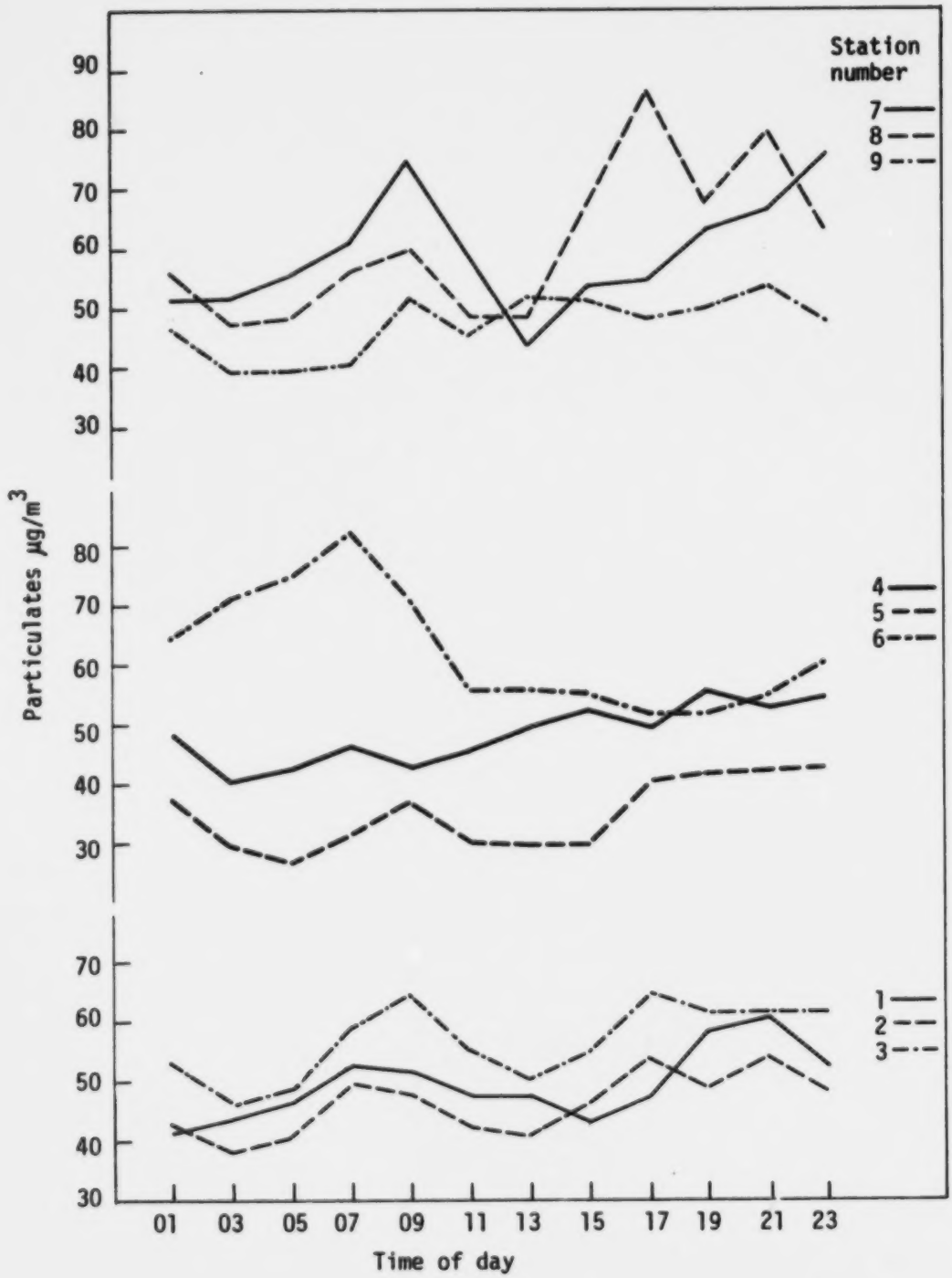

Fig. 16 Average particulate concentration by time of day for indicated stations. A minimum of 10 days of data were averaged at each station.

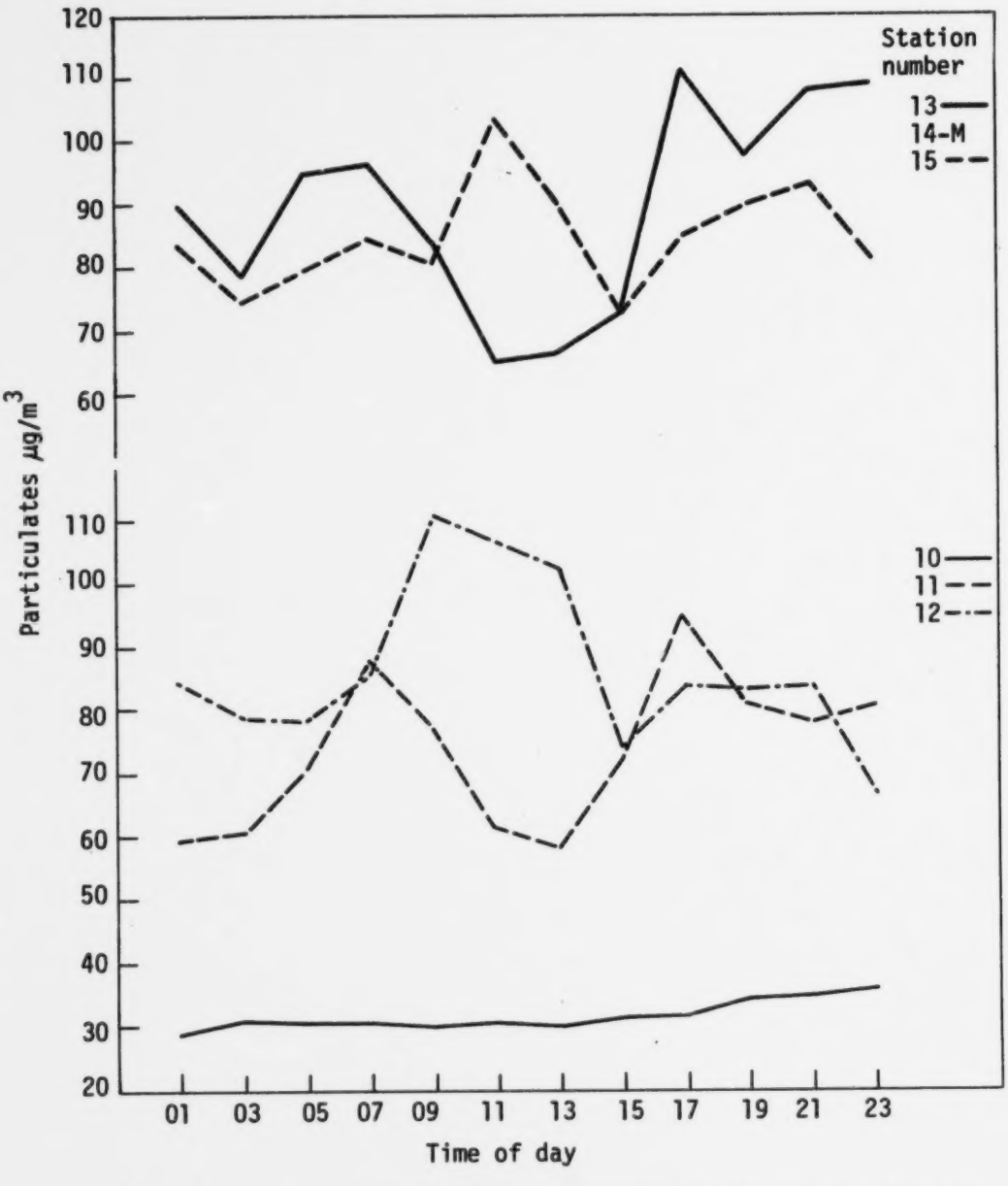

Fig. 17. Average particulate concentration by time of day for indicated stations. A minimum of 10 days of data were averaged at each station.

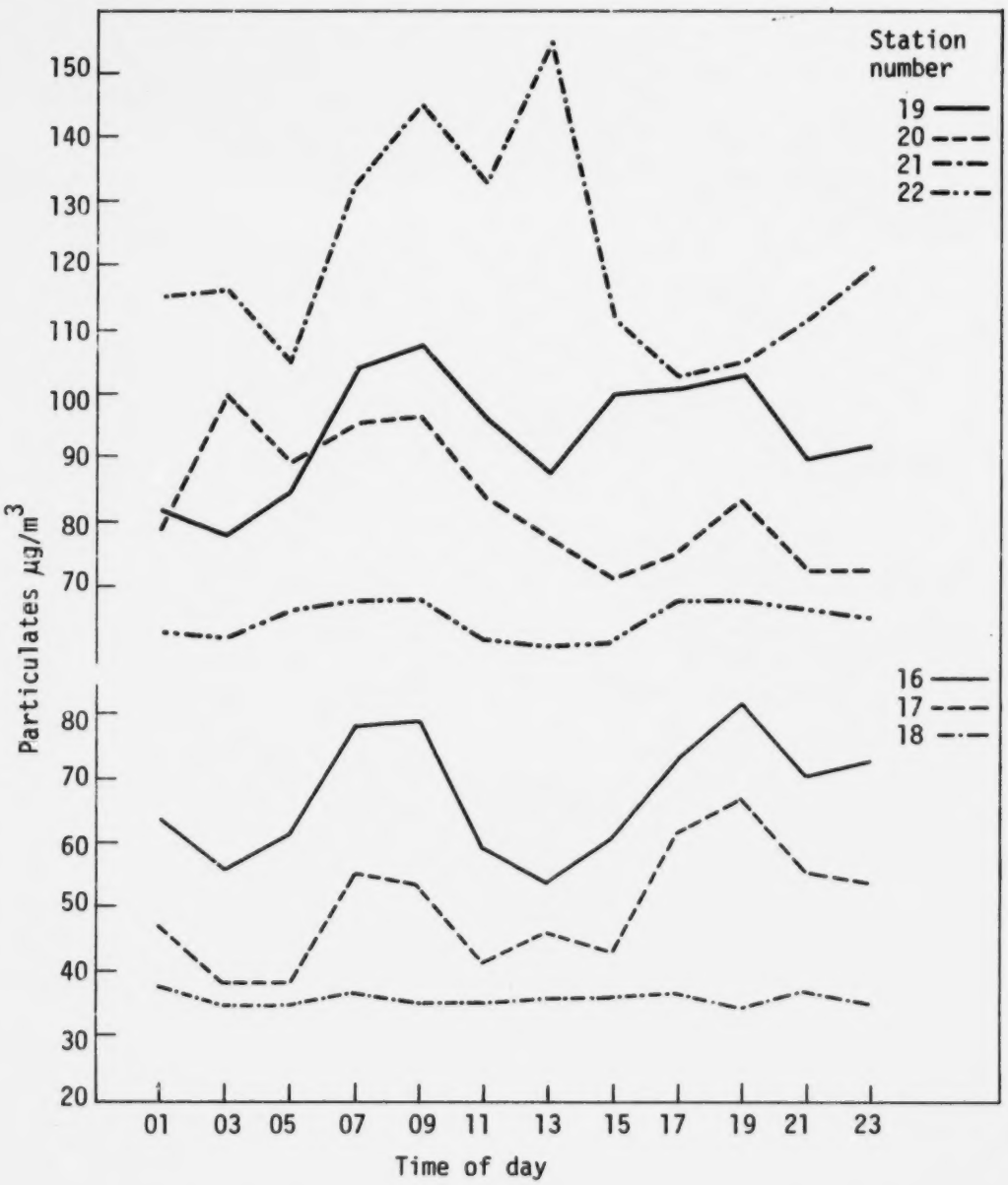

Fig. 18. Average particulate concentration by time of day for indicated stations. A minimum of 10 days of data were averaged at each station.

5.0 SPATIAL VARIABILITY FOR SO_2

5.1 Synoptic Distribution

Mobile surface samplers obtained data that essentially provide a "snap-shot" view of concentration over a route. Three different types of these records are illustrated to show the different character of the measurements.

Figure 19 illustrates successive traverses between Meriden and Greenwich during a test with strong ventilation by a northerly wind. Many of the small spikes are probably caused by moving traffic sources. The major feature, a, is the plume from the Devon power station sampled within three miles of the source. Feature b corresponds to the residential and business area of New Haven.

Figure 20 illustrates three successive traverses from Hartford to New London and records the plumes from two major point sources. At point a the plume from the South Meadows electric generating station is encountered less than one mile from the plant. Feature b is in the position where the plume from the Middletown power plant, 15 miles distant, would be expected. In this case, the rapid variation at low levels is not so evident. Winds were from the southwest and not as strong as for the case in Figure 19.

Figure 21 illustrates a rather unique situation that occurs in Connecticut. This test was conducted with a uniform wind direction of 210 to 220 degrees over the state and the general high level of concentration on the route between Bridgeport and New Milford provide evidence of transport of high SO_2 concentrations from the New York City area. Local sources create rapid changes in concentration even on this generally high background. Feature a is regarded as the edge of the broad plume originating in the New York City

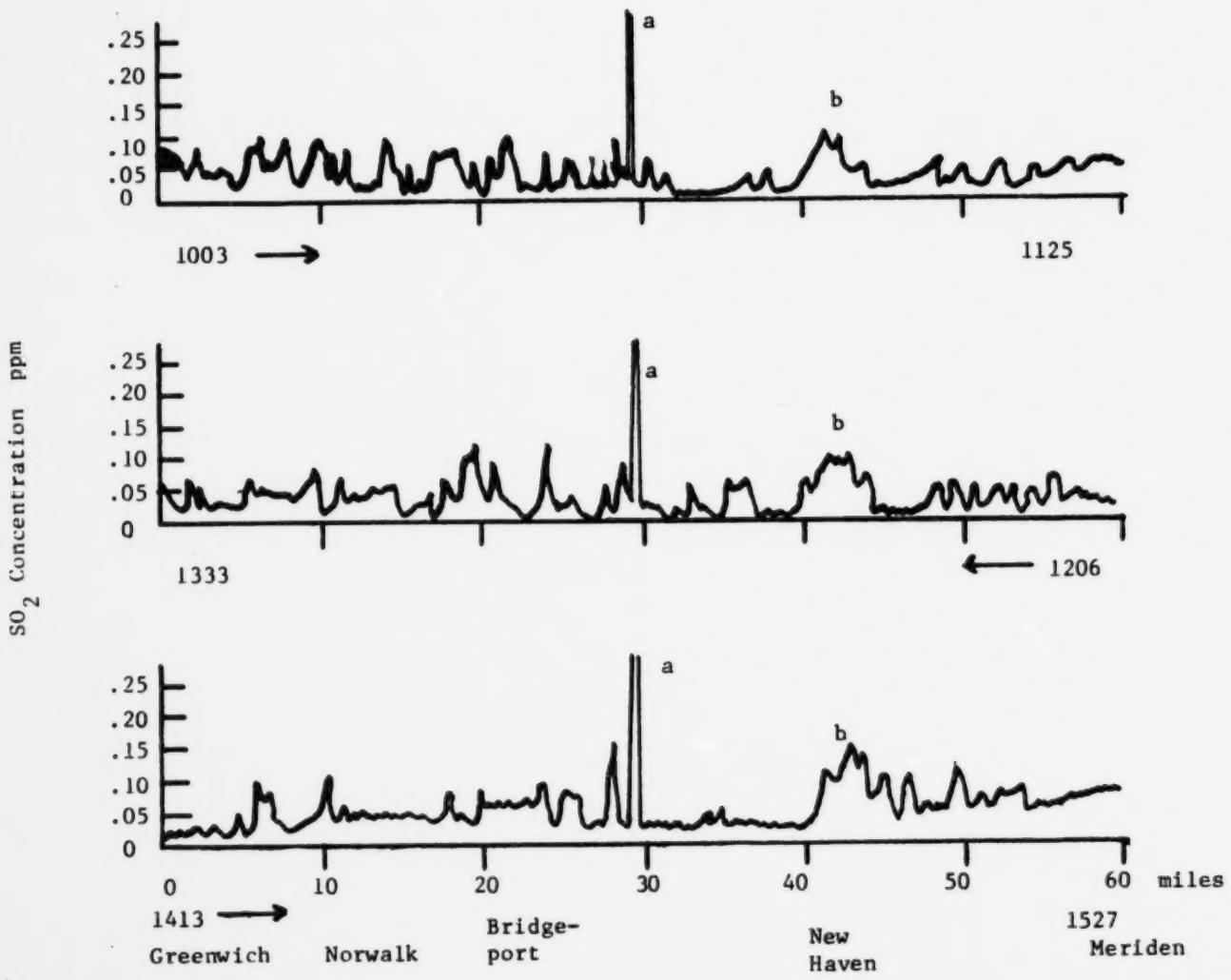

Fig. 19 Test 22 13 March 1969. SO_2 concentrations from automobile traverses between Meriden and Greenwich, Conn. Time of day and direction of travel indicated for each plot.

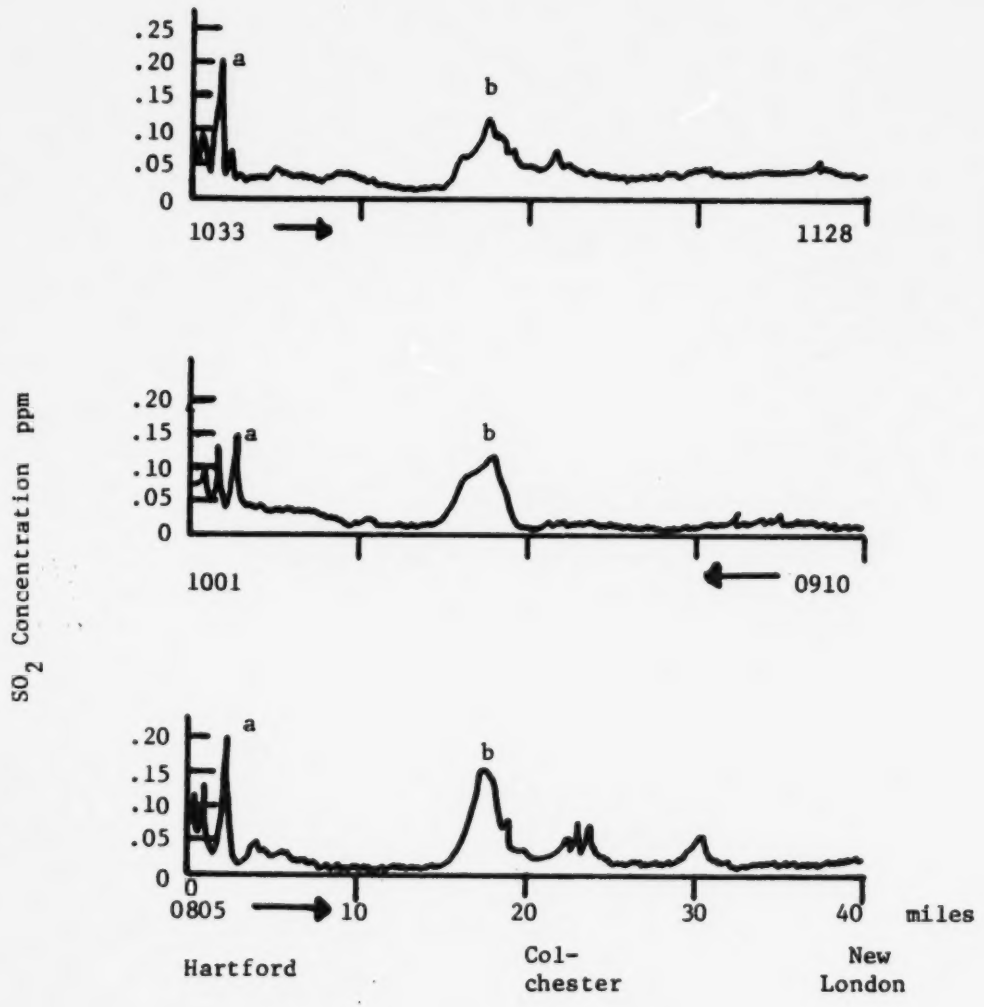

Fig. 20 Test 15 18 December 1968. SO_2 concentrations from automobile traverses between Hartford and New London. Time of day and direction of travel indicated for each plot.

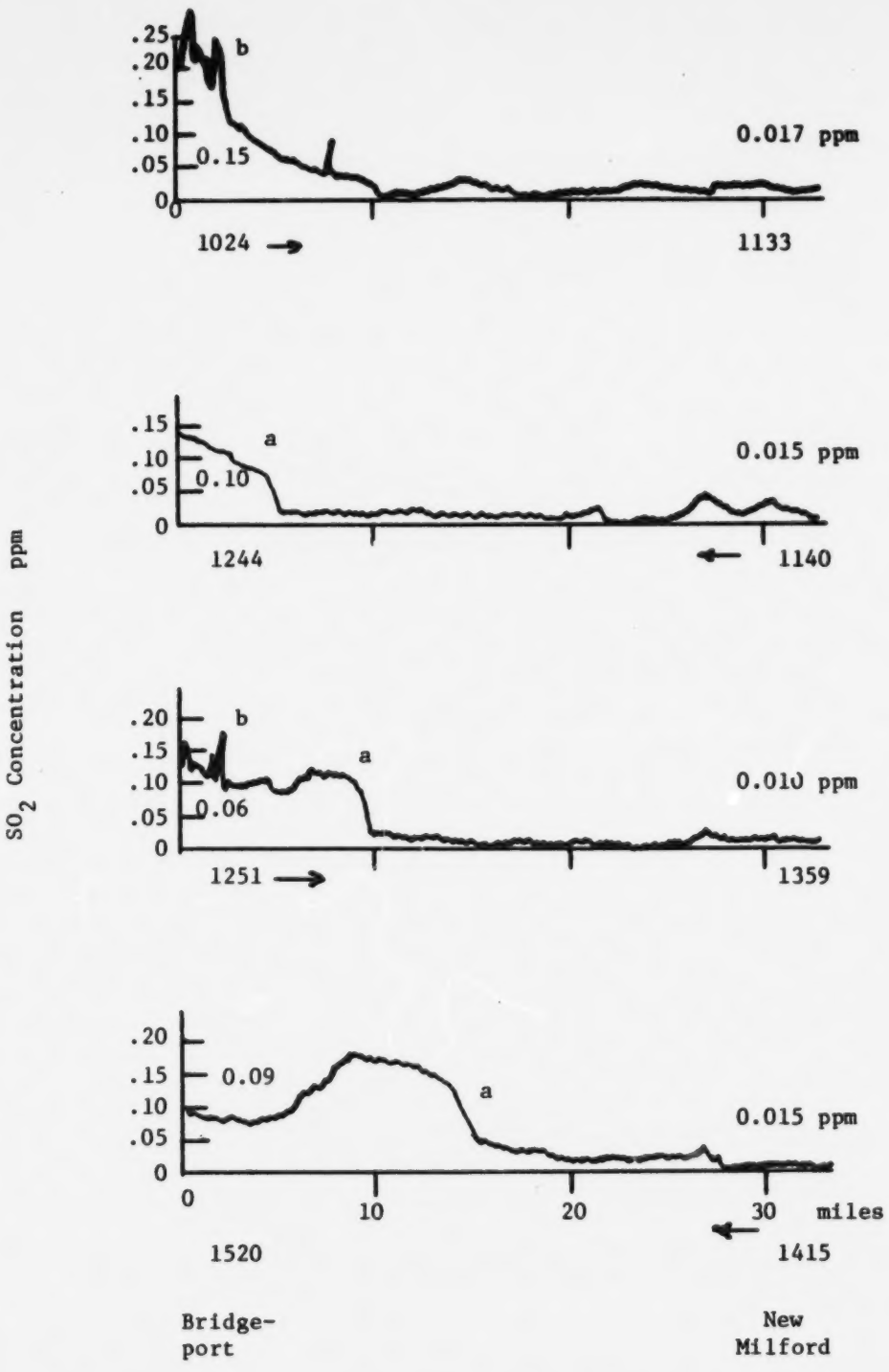

Fig. 21 Test 14 12 December 1968. SO_2 concentration from automobile traverses between Bridgeport and New Milford - Fixed station concentration data for Bridgeport and New Milford is shown.

area, the recurrent feature b is a local source close to the highway. Observed concentrations at Bridgeport and New Milford interpolated from two-hour means are noted on the diagram. They agree within the limits of calibration uncertainty with the constant indications of the mobile sampler.

Certainly a significant feature of all these diagrams is the large variability of concentration in short distances. Ignoring plumes, variability as high as .05 ppm is found in a distance of less than a mile rather frequently in the urban areas and strong wind conditions of Figure 19.

The rural traverse, under wind speeds of about 10 mph, illustrated in Figure 20, indicates that the atmosphere can smooth out the fluctuations given sufficient mixing time. The feature b thought to be caused by the Middletown area some 11 miles away has a base width of two to three miles. Such a well-defined feature is a rather rare occurence.

Figure 21 shows local differences superimposed on a generally high background that is created by sources up to 50 miles distant. In this case it is obvious that modeling must account for the distant sources as well as the local sources even if the distant sources are outside the modeled area. As expected, the great distance has modified the space gradient of concentration so it is relatively flat.

5.2 Model Calculated Variability of SO_2 Concentrations

Nine contiguous 5000 feet grid squares were chosen for examination of grid by grid differences of SO_2 concentrations in four different parts of the state. The centers of the square were located at Bridgeport, New Britain, Norwich and Chaplin. Bridgeport and New Britain are urban locations, Norwich was considered suburban, and Chaplin was a rural area.

Calculations were made with the Martin-Tikvart model (TRW 1969) for the period of the test and the TRC Regional model. Data obtained by averaging all TRC estimates for each two-hour period are shown in Figure 22 with the AQDM estimates and the observed data. Statistics of the TRC estimates are shown in Figure 23.

AQDM estimates were calculated from source strengths averaged over the test days from the TRC model inventory for the areas a minimum of 20 grid squares in all directions from the centers. Observed data were entered for all sites within the data area for each section to provide the model an opportunity to calibrate. Considering the four regions as a whole, no calibration was required. Except for the rural Chaplin area, the AQDM model results in higher estimates, occasionally in reverse gradients from the TRC model and in overestimates in the urban areas where the TRC model resulted in significant underprediction.

This discrepancy in model estimates is puzzling because both are Gaussian in nature and utilize the same source inventory and diffusion curves. The AQDM may have overpredicted because only a neutral stability was used but this seems unlikely to explain the large differences between model calculations.

A second reason may be the arrangement of sources and the observed wind directions. Hilst (1970) showed that the TRC model was especially sensitive to errors in wind direction or lateral placement of the source. Four 5000-ft. grid areas were combined to create the area source terms for the AQDM calculations and some of the significant small scale features could be lost.

Finally, the method of handling vertical diffusion in the source grid-square is slightly different but the methods employed should result in concentration values higher for the short-term model than the long-term model.

Bridgeport			New Britain		
63	97	103	109	145	190
29	39	37	50	47	44
77	118	105	131	185	188
32	(94)	41	127	(139)	56
	35			73	
84	118	117	148	215	195
46	54	62	121	88	52

Norwich			Chaplin		
28	39	41	10	23	31
18	17	22	10	10	11
28	50	48	10	17	22
20	(90)	20	12	(27)	16
	25			10	
45	60	58	9	13	16
22	29	19	13	11	16

Legend: XX AQDM XX TRC (XX) Observed

Fig. 22 Average concentration estimates over all test data collections for SO_2 in contiguous grid squares $\mu g/m^3$.

	Bridgeport			New Britain		
Mean	29	39	37	50	47	44
Stand. Dev.	45	76	64	100	88	64
Mean	32	35	41	127	73	56
Stand. Dev.	52	59	86	99	95	79
Mean	46	54	62	121	88	52
Stand. Dev.	75	105	117	87	85	47

	Norwich			Chaplin		
Mean	18	17	22	10	10	11
Stand. Dev.	24	24	25	7	9	10
Mean	20	25	20	12	10	16
Stand. Dev.	26	35	22	14	6	34
Mean	22	29	19	13	11	16
Stand. Dev.	29	30	19	24	14	30

Fig. 23 TRC Model means and standard deviations of SO_2 concentrations for areas shown in Fig. 22. Concentration of SO_2 in $\mu g/m^3$.

In suburban and rural areas the discrepancies between models is not as large and both underpredict.

Figures 24 to 31 illustrate the cumulative frequency distribution of the observed data, the estimated value for the center square of the 9-square array and the average value of the 9 concentration estimates. Acknowledging that estimates are not good representations of observed data, there is other information available in these figures. For example, the cumulative frequency distribution of the separate grid squares shows that sources were distributed similarly around Bridgeport but not around New Britain. Grid square 2, to the left of the center square at New Britain, had a frequency distribution of calculated concentration very close to that observed in the central grid square. Grid square 1, lower left corner of the area, was nearly identical to 2 and both were significantly different from both the central square and the 9 square average. Obviously the model is capable of generating quite different distributions over a short space scale (5000 ft.) but one does not know a priori which distribution is correct.

Bridgeport observations were obtained in a downtown commercial area. The observed data shown in Figure 24 are nearly log-normal. The slope yields a standard geometric deviation of 2.22, close to CAMP data (Larsen, 1969). There is no significant difference between the estimated center point and the average of the nine grid points. The critical values of difference for the Kolmogorov-Smirnov test are 10.1% for the 0.10 significance level, 11.2% for the 0.05 significance level and 13.6% for the 0.01 significance level. Figure 25 shows the departure of each square from the center grid point (line). There is a significant difference between some of the individual grid point estimates, e.g., position 1 and 6. (The numbering system used is shown on Figure 25.) Some squares depart significantly from

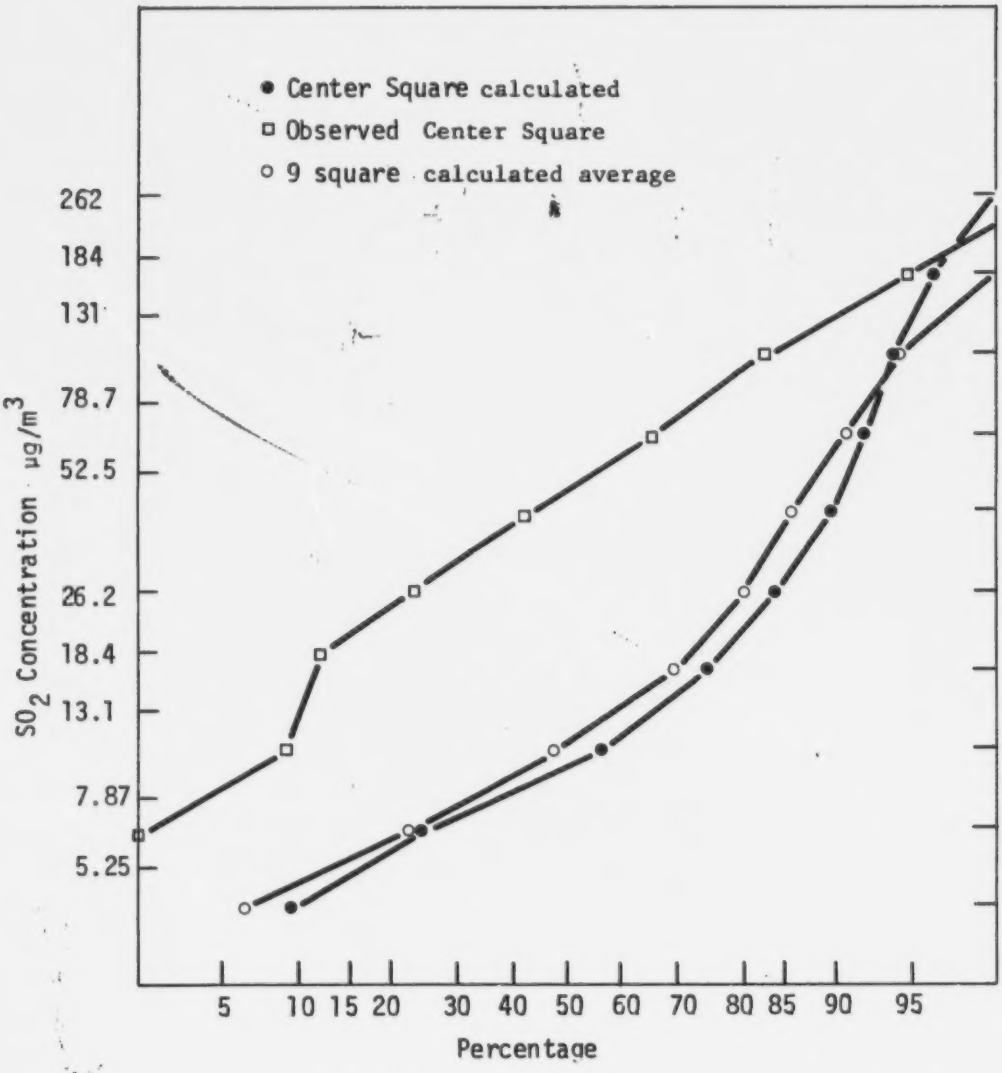

Fig. 24. Bridgeport.

Cumulative frequency distributions of sulfur dioxide concentrations for observed data, center square of interest area and 9 square average of calculated data.

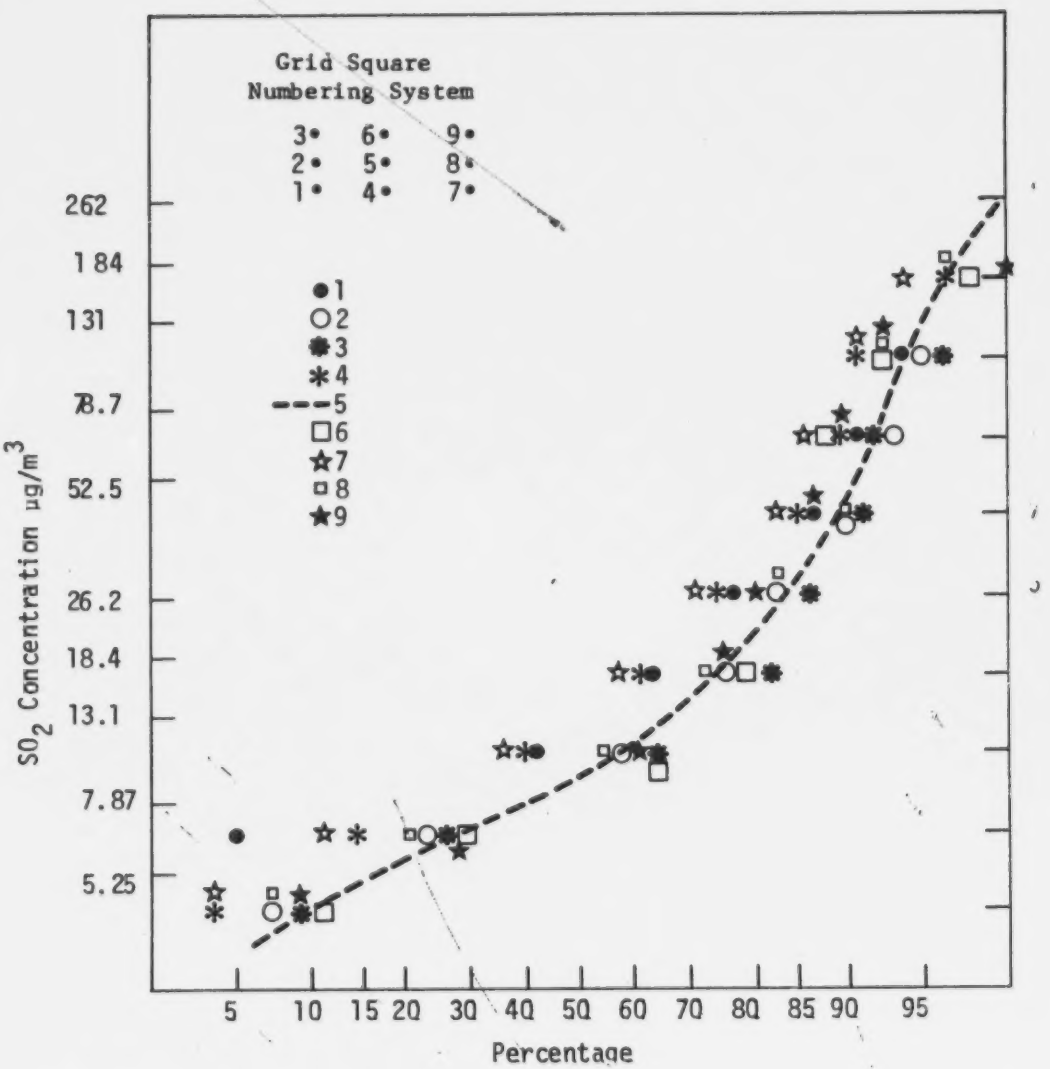

Fig. 25. Bridgeport.

Cumulative frequency distribution of sulfur dioxide concentrations calculated for each of 9 contiguous grid squares.

the average. These differences imply a gradient in the area surrounding the core location.

New Britain data are shown in Figures 26 and 27. The observed data are nearly log-normal with a median value of 77 $\mu g/m^3$ and a geometric standard deviation of 1.85. The slope is quite close to that derived from CAMP data by Larsen. In this case, the 9-square average is low but the shape of the curve is very similar to that of the observed data indicating the source inventory may be too low. The central grid point distribution exhibits the same shape as many of the other locations. The scatter about the center point is clearly indicated in Figure 27. It is obvious that different grid points have significantly different distributions here. Figure 3 indicates that this is a strong source area and the influence is felt by the model. The gradients of the source field are reflected in the gradients in the calculated concentrations. At this location, the cumulative frequency distributions have significantly different slopes from the other three locations. It is taken as an indication that the model can produce appropriate frequency distributions.

Norwich data are illustrated in Figures 28 and 29. This station was in a suburban location but was seriously affected by local sources from nearby buildings and its own roof configuration. Even though the magnitude of the observed Norwich data was biased, the cumulative frequency distribution is nearly log-normal. The steep slope yields a geometric standard deviation of about 4. The estimated data, both average and grid unit values exhibit the familiar concave shape with little difference between most of the estimated values. Squares 4 and 6 depart significantly from the mean values at low concentration but all cluster rather closely at the highest estimated levels.

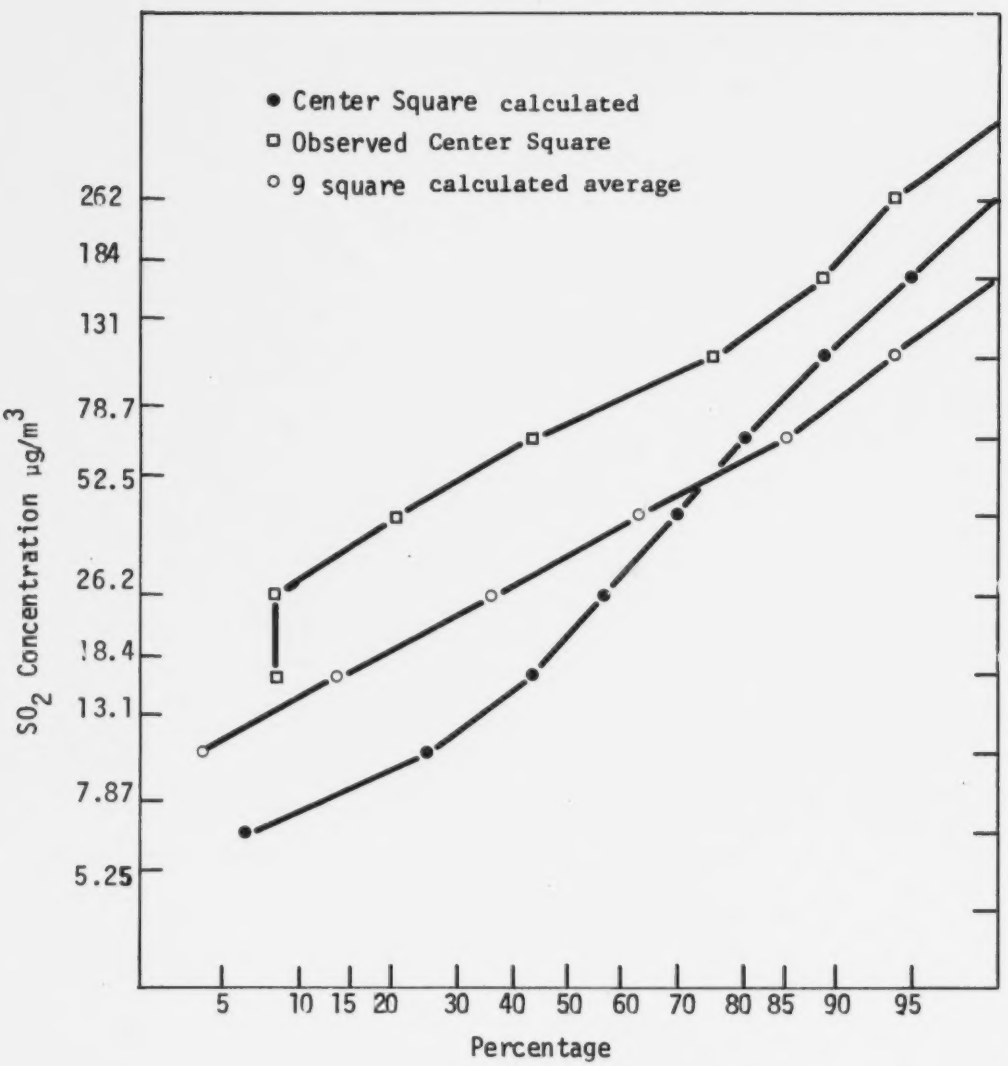

Fig. 26. New Britain.

Cumulative frequency distributions of sulfur dioxide concentrations for observed data, center square of interest area and 9 square average of calculated data.

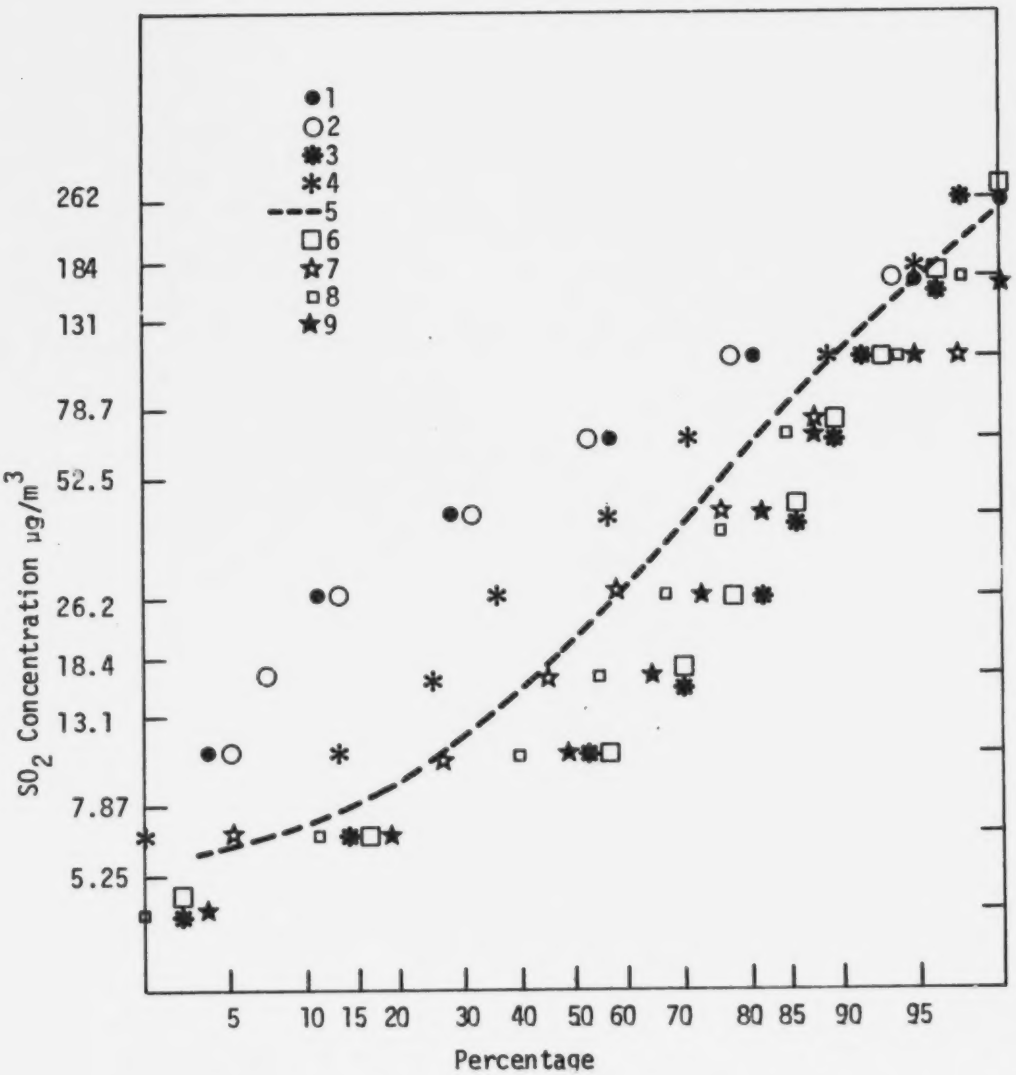

Fig. 27. New Britain.

Cumulative frequency distribution of sulfur dioxide concentrations calculated for each of 9 contiguous grid squares.

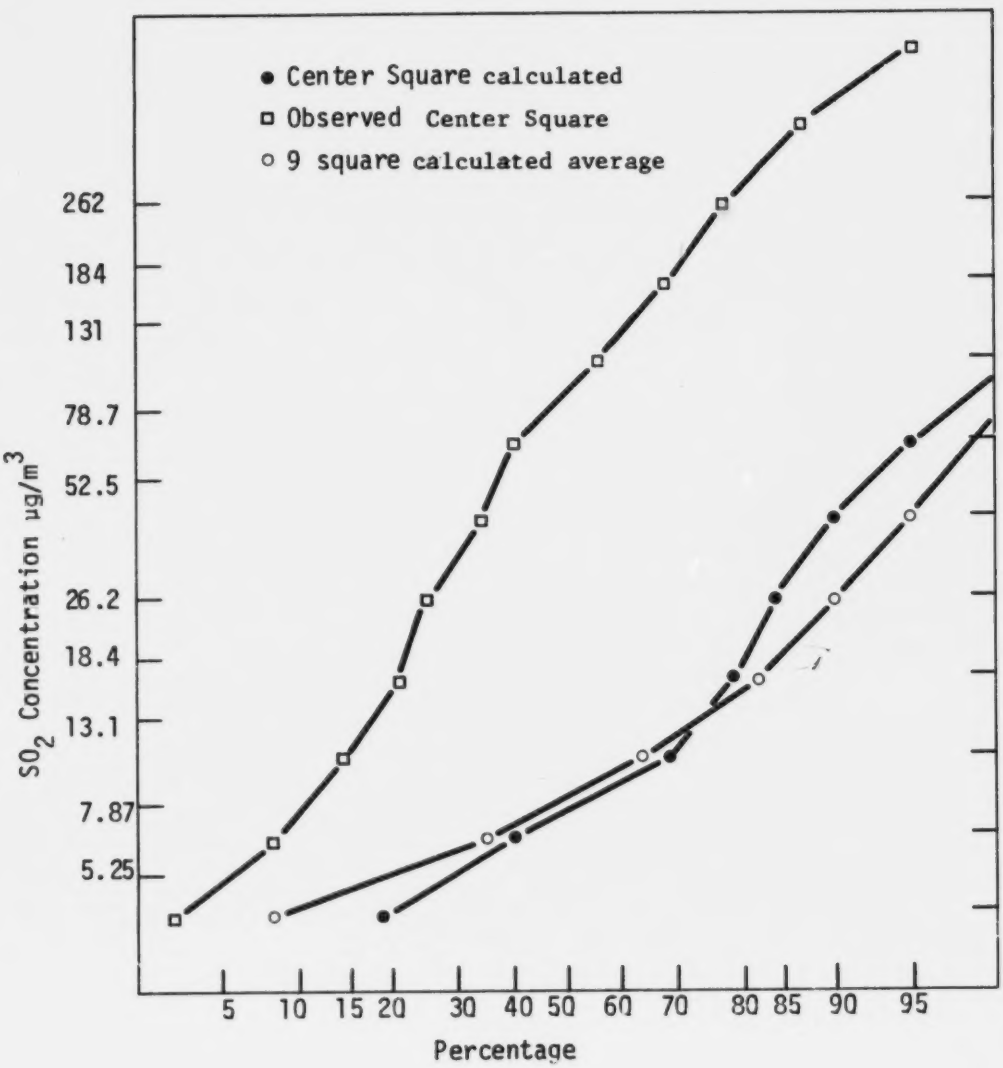

Fig. 28. Norwich.

Cumulative frequency distributions of sulfur dioxide concentrations for observed data, center square of interest area and 9 square average of calculated data.

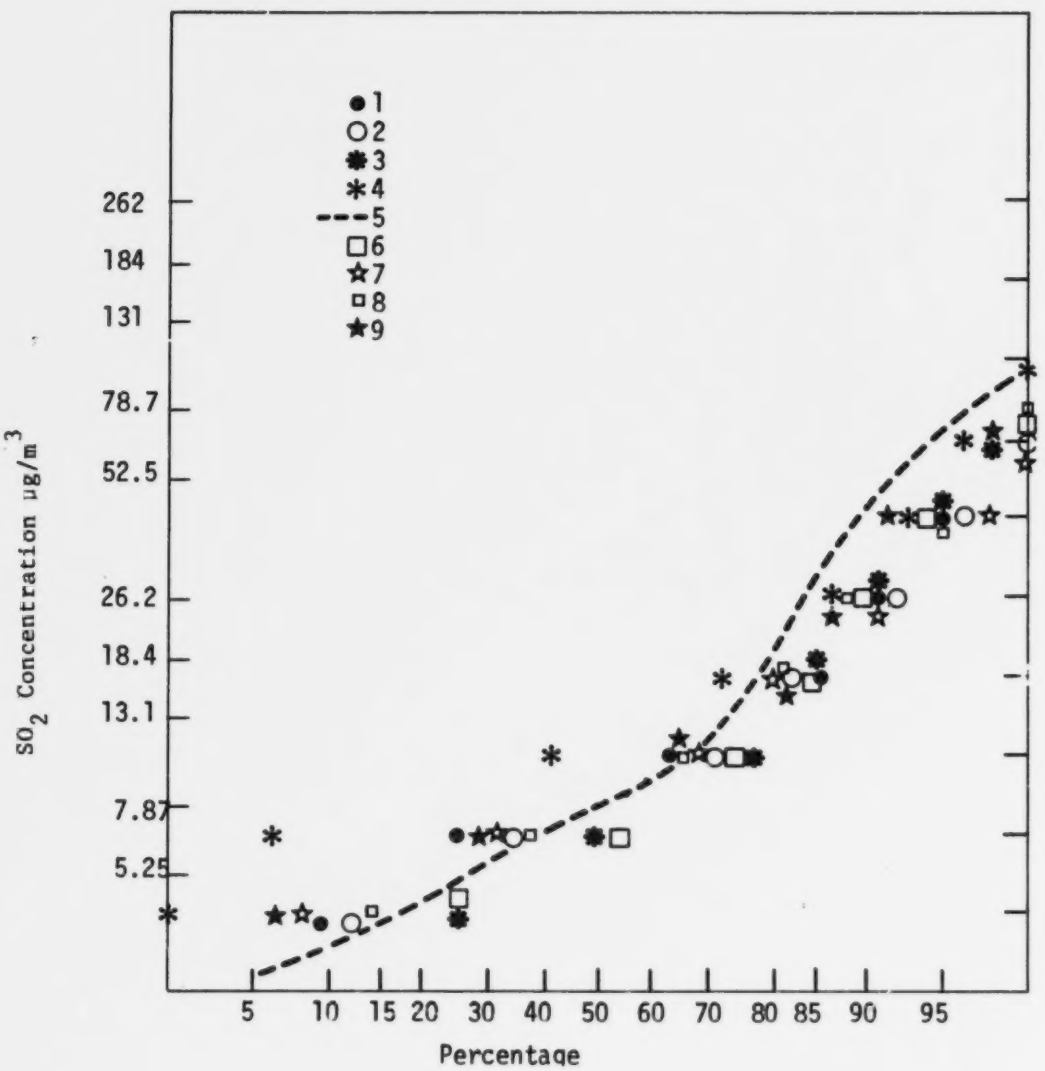

Fig. 29. Norwich

Cumulative frequency distribution of sulfur dioxide concentrations calculated for each of 9 contiguous grid squares.

Chaplin is a rural area with no significant sources in the general area. All of the area source terms are small as shown in Figure 4. The shape of the observed cumulative frequency distribution in Figure 30 comes closer to the concave shape of the estimates than any of the other locations. If a straight line is drawn through the upper 60 percent of the data the low end of the frequency distribution departs significantly from the log-normal distribution. The upper portion has a slope yielding a geometric standard deviation of 2.07, again close to the CAMP data for cities. The estimates show even more concave shapes and again underestimate the observed data by a significant amount. In this case, the differences between grid squares is not significant until the higher concentration estimates are reached. This is the reverse of the usual relationship at other locations.

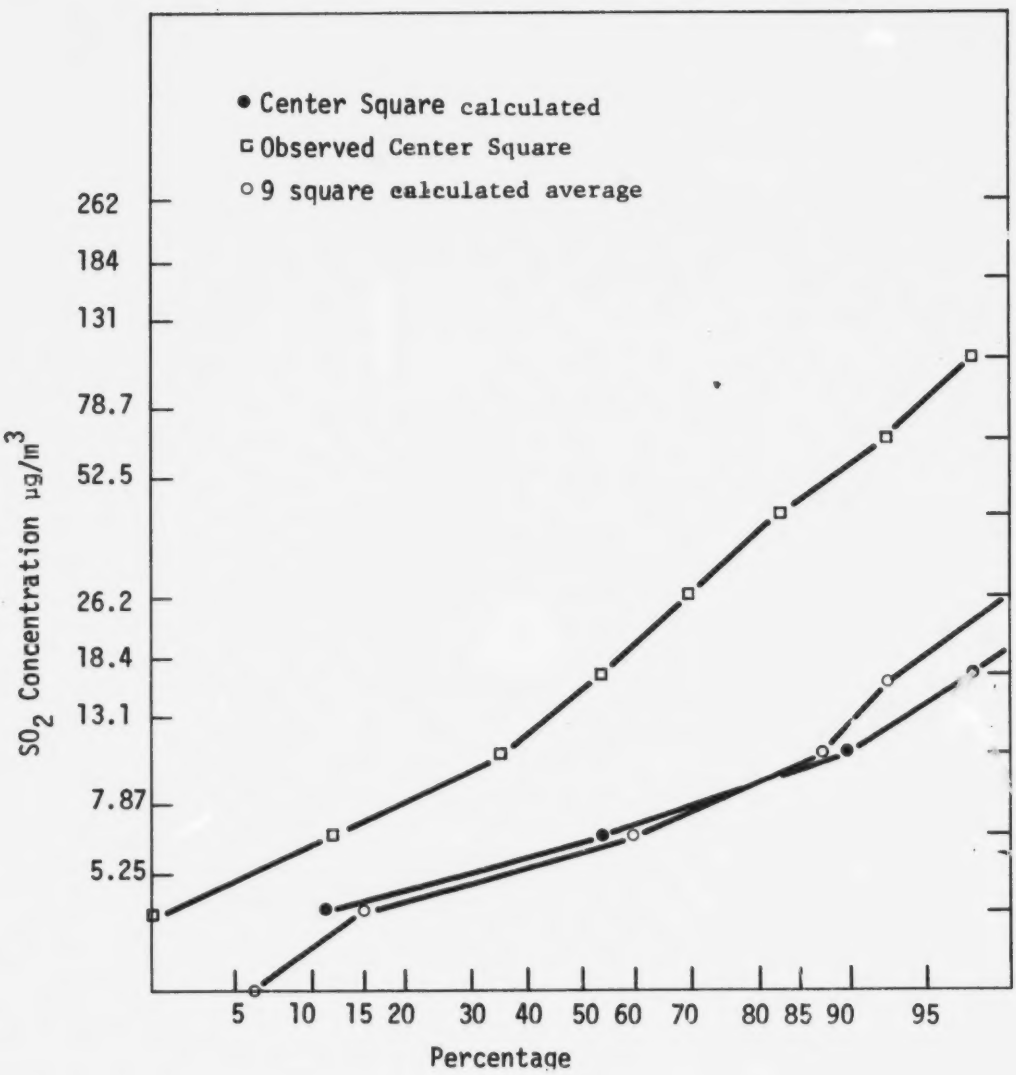

Fig. 30. Chaplin

Cumulative frequency distributions of sulfur dioxide concentrations for observed data, center square of interest area and 9 square average of calculated data.

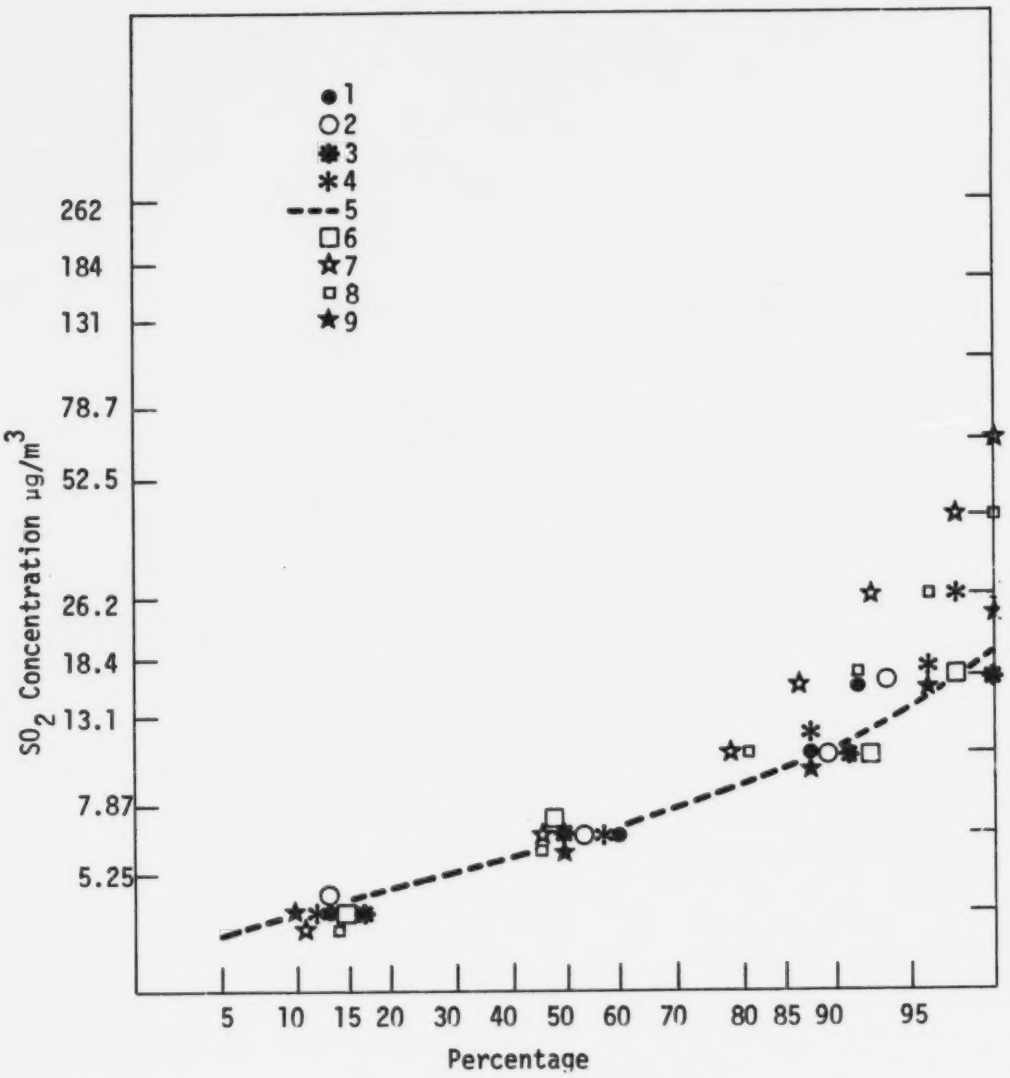

Fig. 31. Chaplin

Cumulative frequency distribution of sulfur dioxide concentrations calculated for each of 9 contiguous grid squares.

6.0 CONCLUSIONS

Analysis of observed 2-hour and 24-hour sulfur dioxide and 24-hour particulate data collected in Connecticut at 30 locations on 25 days during the winter of 1967-68 confirms that a log-normal distribution represents the frequency distribution quite well. Slopes of the log-normal distribution plots were near two, agreeing with Larsen (1969) for urban areas, but steeper slopes were found in suburban and rural areas. The slope of the overall frequency distribution for all data in the state had a slope of three.

Diurnal variation of concentration was consistent with both atmospheric stability changes and emission curves. Peaks coincided with emission peaks but the departure from the mean of the observed concentration data was much less than the relative amplitude of emissions from gas and electric demand curves.

Spatial variability of SO_2 is a function of source pattern, distance from the source, and time period of sampling. In or near urban areas there is a high degree of spatial variability in both instantaneous and short-time averaged air quality measurements. If the sources are diverse enough, the long-term average concentration and frequency distributions may be affected by sampling location. On the basis of data collected in Connecticut, it is concluded that careful selection of a sampling site to minimize local source influence will yield data at a point that is appropriate to an area up to three miles on a side. In areas of strong sources, exemplified here by the New Britain sector, observations at points separated by as little as one mile may have significantly different concentration frequency distributions. If statistics are to be generated for seasonal or annual periods, local gradients will be decreased by the atmospheric averaging process and fewer stations will be required.

It is quite obvious the TRC regional model estimated two-hour concentrations of SO_2 that were too low by a factor of two when compared to observed data. In addition, the shape of the cumulative frequency distribution of estimated values did not agree with observed data, which were usually quite close to log-normal.

The Air Quality Display Model, using the same source inventory produced concentration estimates, in urban areas larger than observed. In the New Britain section where the source inventory was probably best, the AQDM overestimated the observed by one third while the TRC model underestimated by nearly 100 percent. Averaging nine contiguous grid squares does not change these conclusions. It was particularly puzzling to note the reversed estimate gradients around the New Britain sector. It is known that the source inventory in the Middletown area, southeast of New Britain, was overestimated in the source inventory.

Correlation coefficients of observed and estimated data were significant with both models. Cumulative frequency distributions from the AQDM are forced to fit some preassigned function of the log-normal distribution by the Larsen transform. The TRC model estimates were generally not log-normal. They were concave upward, and if the median had been accurate, estimates would have been too high at both the low and high frequency section of the distribution.

Modeling for both short and long-term concentration estimates can be performed but it is very clear that a good source inventory is required. This is especially important for short-term models.

REFERENCES:

Hilst, G. R., 1969: Verification by Observation of the Application of our Air Pollution Model to the State of Connecticut. Final Report to the Connecticut Research Commission, the Travelers Research Corporation.

Hilst, G. R., 1969: The Sensitivities of Air Quality Predictions to Input Errors and Uncertainties. Proceedings of Symposium on "Multiple Source Urban Diffusion Models", EPA Publication AP-86.

Larsen, R. I., 1969: A new Mathematical Model of Air Pollution Concentration Averaging Time and Frequency. J. Air Poll. Control Assoc., 19, 24-30.

Siegal, S., 1956: Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Book Company, Inc., New York, 312 pp.

TRW, Inc., 1969: Air Quality Display Model, Report to U. S. Department of Health, Education and Welfare, Contract No. PH 22-68-60.